

Bulk and soft-edge universality for singular values of products of Ginibre random matrices

Dang-Zheng Liu^{*} Dong Wang[†] and Lun Zhang[‡]

June 9, 2015

Abstract

It has been shown by Akemann, Ipsen and Kieburg that the squared singular values of products of M rectangular random matrices with independent complex Gaussian entries are distributed according to a determinantal point process with a correlation kernel that admits a representation in terms of Meijer G-functions. We prove the universality of the local statistics of the squared singular values, namely, the bulk universality given by the sine kernel and the edge universality given by the Airy kernel. The proof is based on the asymptotic analysis for the double contour integral representation of the correlation kernel. Our strategy can be generalized to deal with other models of products of random matrices introduced recently and to establish similar universal results. Two more examples are investigated, one is the product of M Ginibre matrices and the inverse of K Ginibre matrices studied by Forrester, and the other one is the product of $M - 1$ Ginibre matrices with one truncated unitary matrix considered by Kuijlaars and Stivigny.

1 Introduction and statement of the main results

1.1 Products of Ginibre matrices

Significant progresses have been achieved recently in the study of products of random matrices, which have important applications in Schrödinger operator theory [16], in statistical physics relating to disordered and chaotic dynamical systems [20] and in wireless communication like MIMO (multiple-input and multiple-output) networks [48]. Although the pioneering work of Furstenberg and Kesten [26] focused on the statistical behavior of individual entries in the product as the number of factors tends to infinity, the recent interest of study lies in the distribution of eigenvalues and singular values of the product of a fixed number of matrices, where the sizes of the matrices tend to infinity. Various methods have been applied to perform the spectral analysis in different regimes. Particularly, the tools from free probability allow one to find the limiting mean eigenvalue distributions as in [8, 10, 14, 17, 18, 26, 27, 28, 42, 44, 45]. It turns out that, as in the theory of matrix model for a single random matrix, the various limits exhibit a rich and interesting mathematical structure. Most of the results in literature on the model of products of matrices are about the global spectral properties, but local universality is also

^{*}Key Laboratory of Wu Wen-Tsun Mathematics, Chinese Academy of Sciences, School of Mathematical Sciences, University of Science and Technology of China, Hefei 230026, P. R. China. E-mail: dzliu@ustc.edu.cn

[†]Department of Mathematics, National University of Singapore, 119076, Singapore. E-mail: matwd@nus.edu.sg

[‡]School of Mathematical Sciences and Shanghai Key Laboratory for Contemporary Applied Mathematics, Fudan University, Shanghai 200433, P. R. China. E-mail: lunzhang@fudan.edu.cn

suggested (cf. [7, 22, 23]). Our work is motivated by the previous results and proves the local universality of the squared singular values.

In this paper, we consider $M \geq 1$ independent complex random matrices X_j , $j = 1, \dots, M$, each has size $N_j \times N_{j-1}$ with independent and identically distributed standard complex Gaussian entries. These matrices are also known as complex Ginibre random matrices. We then form the product

$$Y_M = X_M X_{M-1} \cdots X_1. \quad (1.1)$$

For convenience, we assume that N_1, \dots, N_M are associated to a large integer parameter n , which we interchangeably denote by N_0 , such that

$$\min\{N_0, \dots, N_M\} = N_0 = n, \quad (1.2)$$

and set

$$\nu_j = N_j - N_0, \quad j = 0, \dots, M. \quad (1.3)$$

Clearly, $\nu_0 = 0$ and $\nu_j \geq 0$ for $j = 1, \dots, M$.

When $M = 1$, $Y_1 = X_1$ defines the complex Wishart random matrix and plays a fundamental role in random matrix theory; cf. [21]. It is well known that the eigenvalues and squared singular values of Y_1 form determinantal point processes [33, 46], and their distributions are expressed in terms of the correlation kernels. Recent studies find the determinantal structures for the model with general M ; see [3] for the eigenvalues and [6, 7] for the squared singular values. Moreover, further investigations reveal that similar determinantal structures also appear in many other models of products of random matrices, such as the products involving inverses of complex Ginibre matrices [1, 22, 35] and products involving truncated unitary matrices [4, 30, 37]; see also Section 4 below, and the recent review paper [5] and references therein.

We will focus on the squared singular values of Y_M . According to [6], the joint probability density function of the squared singular values is given by (see [6, formula (18)])

$$P(x_1, \dots, x_n) = \frac{1}{Z_n} \prod_{j < k} (x_k - x_j) \det[w_{k-1}(x_j)]_{j,k=1,\dots,n}, \quad (1.4)$$

where $x_j > 0$, the function w_k is a Meijer G-function (see e.g. [11, 40])

$$w_k(x) = G_{0,M}^{M,0} \left(\begin{matrix} - \\ \nu_M, \nu_{M-1}, \dots, \nu_2, \nu_1 + k \end{matrix} \middle| x \right), \quad (1.5)$$

and the normalization constant (see [6, formula (21)]) is

$$Z_n = n! \prod_{i=1}^n \prod_{j=0}^M \Gamma(i + \nu_j).$$

Note that the Meijer G-function $w_k(x)$ can be written as a Mellin-Barnes integral

$$w_k(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \Gamma(s + \nu_1 + k) \prod_{j=2}^M \Gamma(s + \nu_j) x^{-s} ds, \quad k = 0, 1, \dots, \quad (1.6)$$

with $c > 0$.

1.2 The correlation kernel and double integral representation

The determinantal point process (1.4) is a biorthogonal ensemble [15] with correlation kernel

$$K_n(x, y) = \sum_{j=0}^{n-1} \sum_{k=0}^{n-1} x^j (M_n^{-1})_{k,j} w_k(y), \quad (1.7)$$

where M_n is the $n \times n$ matrix of moments

$$M_n = \left(\int_0^\infty x^j w_k(x) dx \right)_{j,k=0,\dots,n-1}. \quad (1.8)$$

Alternatively, one can write the correlation kernel as

$$K_n(x, y) = \sum_{k=0}^{n-1} P_k(x) Q_k(y), \quad (1.9)$$

where for each $k = 0, 1, \dots, n-1$, P_k is a monic polynomial of degree k and Q_k is a linear combination of w_0, \dots, w_k , uniquely defined by the orthogonality

$$\int_0^\infty P_j(x) Q_k(x) dx = \delta_{j,k}. \quad (1.10)$$

Thus the functions P_k and Q_k are the so-called biorthogonal functions. The polynomials P_k are also characterized as multiple orthogonal polynomials [31] with respect to the first M weight functions w_j , $j = 0, \dots, M-1$, as shown in [38].

It turns out that P_k and Q_k are Meijer G-functions [7], and then they have contour integral representations. Then it is shown in [38, Proposition 5.1] that the correlation kernel admits the following double contour integral representation

$$K_n(x, y) = \frac{1}{(2\pi i)^2} \int_{-1/2-i\infty}^{-1/2+i\infty} ds \oint_{\Sigma} dt \prod_{j=0}^M \frac{\Gamma(s + \nu_j + 1)}{\Gamma(t + \nu_j + 1)} \frac{\Gamma(t - n + 1)}{\Gamma(s - n + 1)} \frac{x^t y^{-s-1}}{s - t}, \quad (1.11)$$

where Σ is a closed contour going around $0, 1, \dots, n-1$ in the positive direction and $\operatorname{Re} t > -1/2$ for $t \in \Sigma$. The choices of these two contours are not unique. We can, and indeed will, make some deformations in our later analysis.

1.3 Limiting mean density

The first step of the study of the correlation kernel is to compute the 1-point correlation function, which is also known as the mean density of the model. This global result is also the basis of our proof of the local universality [36]. As mentioned at the very beginning, the limiting mean density/spectral distribution of the squared singular values for Y_M is well understood using tools from free probability; see also recent work [41, 49] for the study from the polynomials P_k . It turns out that, after proper scaling, the limiting mean density is recognized as the Fuss-Catalan distribution [8, 10, 42], i.e., its k -th moment is given by the Fuss-Catalan number

$$\frac{1}{Mk+1} \binom{Mk+k}{k}, \quad k = 0, 1, 2, \dots \quad (1.12)$$

The (rescaled) limiting mean density is supported on an interval $[0, (M+1)^{M+1}/M^M]$, with explicit form given in terms of Meijer G-functions [45] or multivariate integrals [39]. Probably the

simplest form of the density function for general M is expressed by the following parametrization of the argument [14, 29, 41]:

$$x = \frac{(\sin((M+1)\varphi))^{M+1}}{\sin \varphi (\sin(M\varphi))^M}, \quad 0 < \varphi < \frac{\pi}{M+1}. \quad (1.13)$$

It is readily seen that this parametrization is a strictly decreasing function of φ , thus gives a one-to-one mapping from $(0, \pi/(M+1))$ to $(0, (M+1)^{M+1}/M^M)$. The density function in terms of φ is then given by

$$\rho(\varphi) = \frac{1}{\pi x} \frac{\sin((M+1)\varphi)}{\sin(M\varphi)} \sin \varphi = \frac{1}{\pi} \frac{(\sin \varphi)^2 (\sin(M\varphi))^{M-1}}{(\sin((M+1)\varphi))^M}. \quad (1.14)$$

From (1.13) and (1.14), one can check (cf. [23]) that the density blows up with a rate $x^{-M/(M+1)}$ near the origin (hard edge), while vanishes as a square root near $(M+1)^{M+1}/M^M$ (soft edge). These facts particularly suggest, as pointed out in [7], the classical bulk and soft edge universality (via the sine kernel and Airy kernel, respectively) should hold in the bulk and the right edge respectively as in the $M = 1$ case, that is, the complex Wishart ensemble, but new limiting distributions are required to describe the local behavior at the hard edge if $M > 1$. The new limiting distributions, characterized by their limiting correlation kernels, were computed in [38] by taking limit from the integral representation (1.11). Here we note that the new family of kernels is a generalization of the classical Bessel kernel [47] which is the $M = 1$ case of the family, and they are universal correlation kernels since they also appear in many other random models, including Cauchy-chain matrix models [12, 13], products of Ginibre matrices with inverse ones [22], biorthogonal ensembles of Borodin [15] (as shown in [37]), etc. However, the conceptually simpler universality results in the bulk and at the right edge turn out to be technically more complicated and have been left open in [38].

It is the aim of this paper to confirm the bulk and soft edge universality in the products of Ginibre matrices Y_M . Our main results are stated in the next section.

1.4 Statement of the main results

We start with the definition of the sine kernel (see [9, Theorem 3.1.1]; here we take a different normalization):

$$K_{\sin}(x, y) := \frac{\sin \pi(x - y)}{\pi(x - y)}. \quad (1.15)$$

Recall the correlation kernel $K_n(x, y)$ given in (1.11), our first result is stated as follows:

Theorem 1.1 (Bulk universality). *For $x_0 \in (0, (M+1)^{M+1}/M^M)$, which is parametrized through (1.13) by $\varphi = \varphi(x_0) \in (0, \pi/(M+1))$, we have, with ν_1, \dots, ν_M being fixed,*

$$\lim_{n \rightarrow \infty} \frac{e^{-\pi \xi \cot \varphi}}{e^{-\pi \eta \cot \varphi}} \frac{n^{M-1}}{\rho(\varphi)} K_n \left(n^M \left(x_0 + \frac{\xi}{n \rho(\varphi)} \right), n^M \left(x_0 + \frac{\eta}{n \rho(\varphi)} \right) \right) = K_{\sin}(\xi, \eta) \quad (1.16)$$

uniformly for ξ and η in any compact subset of \mathbb{R} , where $\rho(\varphi)$ is defined in (1.14).

Next, recall the Airy kernel defined by [9, Section 3.1]

$$K_{\text{Ai}}(x, y) := \frac{\text{Ai}(x)\text{Ai}'(y) - \text{Ai}'(x)\text{Ai}(y)}{x - y} = \frac{1}{(2\pi i)^2} \int_{\gamma_R} d\mu \int_{\gamma_L} d\lambda \frac{e^{\frac{\mu^3}{3} - x\mu}}{e^{\frac{\lambda^3}{3} - y\lambda}} \frac{1}{\mu - \lambda}, \quad (1.17)$$

where γ_R and γ_L are symmetric with respect to the imaginary axis, and γ_R is a contour in the right-half plane going from $e^{-\pi i/3} \cdot \infty$ to $e^{\pi i/3} \cdot \infty$; see Figure 1 for an illustration.

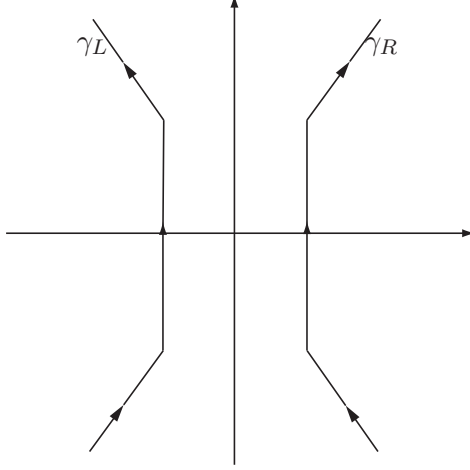


Figure 1: The contours γ_L and γ_R in the definition of Airy kernel.

Theorem 1.2 (Soft edge universality). *With ν_1, \dots, ν_M being fixed, we have*

$$\lim_{n \rightarrow \infty} \frac{e^{-2^{-\frac{1}{3}}(M+1)^{\frac{2}{3}}\xi n^{\frac{1}{3}}}}{e^{-2^{-\frac{1}{3}}(M+1)^{\frac{2}{3}}\eta n^{\frac{1}{3}}}} n^{M-\frac{2}{3}} c_2 K_n \left(n^M \left(x_* + \frac{c_2 \xi}{n^{\frac{2}{3}}} \right), n^M \left(x_* + \frac{c_2 \eta}{n^{\frac{2}{3}}} \right) \right) = K_{\text{Ai}}(\xi, \eta) \quad (1.18)$$

uniformly for ξ and η in any compact subset of \mathbb{R} , where

$$x_* = \frac{(M+1)^{M+1}}{M^M} \quad \text{and} \quad c_2 = \frac{(M+1)^{M+\frac{2}{3}}}{2^{\frac{1}{3}} M^{M-1}}. \quad (1.19)$$

Theorems 1.1 and 1.2 then imply that the universal scaling limits of the correlation kernel (in the bulk or at the soft edge) that are typical for unitary random matrix ensembles also occur in products of complex Ginibre random matrices.

Remark 1. If we strengthen the result in Theorem 1.2 from uniform convergence into the trace norm convergence of the integral operators with respect to the correlation kernels, then as a direct consequence we have that the limiting distribution of the largest squared singular value, after rescaling, converges to the Tracy-Widom distribution [9, Theorem 3.1.5]. Since the proof of trace norm convergence is only a technical elaboration that confirms a well-expected result, we do not give the detail.

1.5 About the proof

Our proof of the main theorems is based on a steepest descent analysis of the double contour integral (1.11), whose integrand contains products and ratios of gamma functions with large arguments. By Stirling's formula, the logarithms of the gamma functions are approximated by elementary functions for n large, which play the role of phase function. In the bulk regime there are two complex conjugate saddle points, while in the edge regime these two saddle points coalesce into one. The main challenge of the proof is to find suitable contours of integration and sophisticated estimates of integrals. As we shall see later, the parametrization (1.13) will be essential in the analysis. Our strategy can be generalized to deal with some other product models introduced recently, where the correlation kernels have similar structures. We will also discuss about this aspect at the end of this paper.

Here we note that the steepest descent analysis of a double contour integral involving gamma functions was used in a different random matrix model in [2], where the limiting Pearcey kernel

was derived. The preprint [25] that considers a random matrix model similar to that in [2] applies the method detailed in this paper to perform asymptotic analysis. See also [50].

The rest of this paper is organized as follows. Theorems 1.1–1.2 are proved in Section 2, upon two technical lemmas that are proved in Section 3 on the properties of the specially deformed integral contours. Section 3 also contains the precise construction of the deformed contours. We conclude this paper in Section 4 with a discussion on the generalizations of our method to establish similar universal results in other models of products of random matrices. We present two more examples, one is the product of M Ginibre matrices and the inverse of K Ginibre matrices studied by Forrester [22], and the other one is the product of $M - 1$ Ginibre matrices with one truncated unitary matrix considered by Kuijlaars and Stivigny [37].

2 Proofs of the main theorems

2.1 Notations and contour deformations

For notational simplicity, we set

$$F(z; a) := \log \left(\frac{\prod_{j=0}^M \Gamma(z + \nu_j + 1)}{\Gamma(z - n + 1)} a^{-z} \right), \quad a \geq 0, \quad (2.1)$$

where the logarithm takes the principal branch and we assume that the value of $\log z$ for $z \in (-\infty, 0)$ is continued from above. The asymptotics of F is crucial in our analysis. To proceed, note the Stirling's formula for gamma function [43, formula 5.11.1] reads

$$\log \Gamma(z) = \left(z - \frac{1}{2} \right) \log z - z + \frac{1}{2} \log(2\pi) + \mathcal{O}\left(\frac{1}{z}\right) \quad (2.2)$$

as $z \rightarrow \infty$ in the sector $|\arg z| \leq \pi - \epsilon$ for some $\epsilon > 0$. Hence it follows that if $|z| \rightarrow \infty$ and $|z - n| \rightarrow \infty$, while $\arg z$ and $\arg(z - n)$ are in $(-\pi + \epsilon, \pi - \epsilon)$, then uniformly

$$\begin{aligned} F(z; a) = \tilde{F}(z; a) + \sum_{j=0}^M \left(\nu_j + \frac{1}{2} \right) \log z - \frac{1}{2} \log(z - n) \\ + \frac{M}{2} \log(2\pi) + \mathcal{O}(\min(|z|, |z - n|)^{-1}), \end{aligned} \quad (2.3)$$

where

$$\tilde{F}(z; a) = (M + 1)z(\log z - 1) - (z - n)(\log(z - n) - 1) - z \log a. \quad (2.4)$$

Furthermore, we have

$$\tilde{F}(nz; n^M a) = n\hat{F}(z; a) + n \log n, \quad (2.5)$$

where

$$\hat{F}(z; a) = (M + 1)z(\log z - 1) - (z - 1)(\log(z - 1) - 1) - z \log a. \quad (2.6)$$

The behaviour of $\operatorname{Re} \hat{F}(z; a)$ is crucial in the saddle point analysis in this paper, and we plot the level line $\operatorname{Re} \hat{F}(z; a) = \operatorname{Re} \hat{F}(w_{\pm}; a)$ in Figure 4, where w_{\pm} are the critical points of $\hat{F}(z; a)$, see (2.10) and (2.11).

To prove the results of universality, we also need to deform the contours in (1.11). First we note that the integral contour for s can be replaced by any infinite contour \mathcal{C} that is taken to go from $-i\infty$ to $i\infty$, as long as Σ is on the right side of \mathcal{C} . Thus (with shorthand notation F defined in (2.1)), we express (1.11) as

$$K_n(x, y) = \frac{y^{-1}}{(2\pi i)^2} \int_{\mathcal{C}} ds \oint_{\Sigma} dt \frac{e^{F(s; y)}}{e^{F(t; x)}} \frac{1}{s - t}. \quad (2.7)$$

In the proof of the soft edge universality, we will further deform \mathcal{C} in (2.7) such that Σ is on its left, and it turns out that the resulting double contour integral remains the same. To see this, let \mathcal{C} and \mathcal{C}' be two infinite contours from $-i\infty$ to $i\infty$ such that Σ lies between \mathcal{C} and \mathcal{C}' , that is, Σ is enclosed by $\mathcal{C} \cup \mathcal{C}'$. Applying the residue theorem to the integral on $\mathcal{C} \cup \mathcal{C}'$, it follows

$$\int_{\mathcal{C}} ds \oint_{\Sigma} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} - \int_{\mathcal{C}'} ds \oint_{\Sigma} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} = 2\pi i \int_{\Sigma} \left(\frac{x}{y}\right)^t dt = 0. \quad (2.8)$$

Hence, the double contour integral does not change if \mathcal{C} is replaced by \mathcal{C}' . The deformation of \mathcal{C} is shown in Figure 2.

Similarly, we can show that if Σ is split into two disjoint closed counterclockwise contours $\Sigma = \Sigma_1 \cup \Sigma_2$, which jointly enclose poles $0, 1, \dots, n-1$, and \mathcal{C} is an infinite contour from $-i\infty$ to $i\infty$ such that Σ_1 is on the left side of \mathcal{C} and Σ_2 is on the right side of \mathcal{C} , the formula (2.7) is still valid. We will use such kind of contours in the proof of the bulk universality. The deformation of Σ is shown in Figure 2.

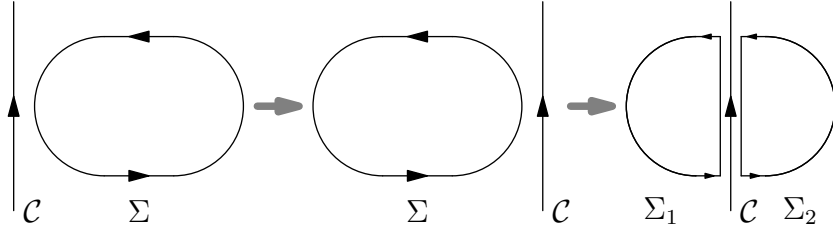


Figure 2: In the contour integral (2.7), the position of Σ and \mathcal{C} can be switched, and Σ can be split into Σ_1 and Σ_2 and \mathcal{C} goes between them.

To facilitate the asymptotic analysis, throughout the rest of this paper, we shall denote by $D_r(a)$ the disc centered at a with radius r , and by \mathbb{C}_{\pm} the upper/lower half complex plane, respectively, and if C is a contour in \mathbb{C} and $r > 0$, then denote by rC the contour $\{z \in \mathbb{C} \mid z/r \in C\}$ with the same orientation as C .

2.2 Proof of Theorem 1.1

Before going in the detail of the proof, we sketch the strategy. As mentioned in the end of Section 2.1, we use the integral representation (2.7) of the kernel $K_n(x, y)$, with the contours deformed, such that Σ is split into two parts, and the contour \mathcal{C} goes between them. But in the asymptotic analysis, we group the “curved” part and the “vertical” part of Σ into Σ_{curved} and Σ_{vertical} respectively; see Figure 3 to get a visual idea of these two parts. Then we compute $K_n(x, y)$ as $I_1 + I_2$, where I_1 is the integral (2.7) over \mathcal{C} and Σ_{curved} , while I_2 is that over \mathcal{C} and Σ_{vertical} . I_1 is evaluated by the usual saddle point method, as detailed in Section 2.2.3, and it turns out to be the insignificant part; I_2 is evaluated by an application of Cauchy’s theorem, see Section 2.2.2, and it turns out to be the main contribution.

2.2.1 Exact deformation of the contours

For any $x_0 \in (0, (M+1)^{M+1}/M^M)$, we use the parametrization (1.13), and let φ be the unique real number in $(0, \pi/(M+1))$ such that

$$x_0 = \frac{(\sin((M+1)\varphi))^{M+1}}{\sin \varphi (\sin(M\varphi))^M}.$$

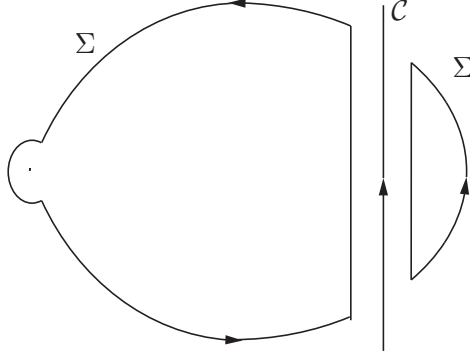


Figure 3: The contours \mathcal{C} and Σ defined in (2.13) and (2.14), respectively.

To prove the bulk universality, we assume that the arguments x and y in (2.7) are in the form

$$x = n^M \left(x_0 + \frac{\xi}{n\rho(\varphi)} \right), \quad y = n^M \left(x_0 + \frac{\eta}{n\rho(\varphi)} \right), \quad (2.9)$$

where ξ and η are in a compact subset of \mathbb{R} and $\rho(\varphi)$ is given in (1.14). For the asymptotic analysis, we denote

$$w_{\pm} = \frac{\sin((M+1)\varphi)}{\sin(M\varphi)} e^{\pm i\varphi}. \quad (2.10)$$

They are two saddle points of the function $\hat{F}(z; x_0)$ defined in (2.6), for

$$\hat{F}_z(w_{\pm}; x_0) := \left. \frac{d}{dz} \hat{F}(z; x_0) \right|_{z=w_{\pm}} = 0. \quad (2.11)$$

It is also straightforward to check

$$\hat{F}_{zz}(w_{\pm}; x_0) := \left. \frac{d^2}{dz^2} \hat{F}(z; x_0) \right|_{z=w_{\pm}} = \frac{1}{w_{\pm}} \left(M+1 - \frac{\sin((M+1)\varphi)}{\sin \varphi} e^{\mp iM\varphi} \right). \quad (2.12)$$

The shapes of the contours \mathcal{C} and Σ in (2.7) used in this subsection are schematically illustrated in Figure 3, and are precisely described as follows, based on the two contours $\tilde{\mathcal{C}}_{x_0}$ and $\tilde{\Sigma}^{\epsilon}$ explicitly constructed in Section 3.1. The contour \mathcal{C} is chosen to be

$$\mathcal{C} = n\tilde{\mathcal{C}}_{x_0}, \quad (2.13)$$

i.e., \mathcal{C} is the vertical, upward contour through the points nw_+ and nw_- . To describe Σ , we let ϵ and ϵ' be two small enough positive constants. Then the contour Σ is defined by

$$\Sigma = \Sigma_{\text{curved}} \cup \Sigma_{\text{vertical}}, \quad (2.14)$$

where

$$\Sigma_{\text{curved}} = \Sigma_1 \cup \Sigma_2, \quad \Sigma_{\text{vertical}} = \Sigma_3 \cup \Sigma_4, \quad (2.15)$$

and

$$\begin{aligned} \Sigma_1 &= n\tilde{\Sigma}^r \cap \{z \mid \operatorname{Re} z \leq \operatorname{Re} nw_{\pm} - \epsilon\}, \\ \Sigma_2 &= n\tilde{\Sigma}^r \cap \{z \mid \operatorname{Re} z \geq \operatorname{Re} nw_{\pm} + \epsilon\}, \\ \Sigma_3 &= \text{vertical bar connecting the two ending points of } \Sigma_1, \\ \Sigma_4 &= \text{vertical bar connecting the two ending points of } \Sigma_2. \end{aligned} \quad \text{with } r = \frac{[\epsilon'n] + \frac{1}{2}}{n}, \quad (2.16)$$

Note that Σ consists of two disjoint closed contours: $\Sigma_1 \cup \Sigma_3$ and $\Sigma_2 \cup \Sigma_4$, whose orientations are counterclockwise. By the arguments at the end of Section 2.1, such kind of contour deformation is allowed. Here we assume that $\operatorname{Re} nw_{\pm}$ is not an integer, so that \mathcal{C} does not pass through any integer point. We further assume that ϵ is small enough so that \mathcal{C} , Σ_3 and Σ_4 all lie between two consecutive integers k and $k+1$. In the case that $\operatorname{Re} nw_{\pm} \in \mathbb{Z}$ and \mathcal{C} passes through an integer point, we simply shift contour \mathcal{C} horizontally by $1/2$ to make it go between two consecutive integers, and all arguments below work in the same way.

In the asymptotic analysis of $K_n(x, y)$, we write (2.7) as

$$K_n(x, y) = I_1 + I_2, \quad \text{where} \quad \begin{Bmatrix} I_1 \\ I_2 \end{Bmatrix} = \frac{y^{-1}}{(2\pi i)^2} \times \begin{Bmatrix} \int_{\mathcal{C}} ds \oint_{\Sigma_{\text{curved}}} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t}, \\ \int_{\mathcal{C}} ds \oint_{\Sigma_{\text{vertical}}} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t}. \end{Bmatrix} \quad (2.17)$$

The following properties of $F(z; n^M x_0)$ on the contours Σ_{curved} and \mathcal{C} will play an important role in our later analysis.

Lemma 2.1. *There exists a constant $\delta > 0$ such that for n large enough,*

$$\operatorname{Re} F(z; n^M x_0) \geq \operatorname{Re} F(nw_{\pm}; n^M x_0) + \delta n \left| \frac{z}{n} - w_{\pm} \right|^2 \quad \text{for } z \in \Sigma_{\text{curved}} \cap D_{n^{\frac{3}{5}}}(nw_{\pm}), \quad (2.18)$$

$$\operatorname{Re} F(z; n^M x_0) > \operatorname{Re} F(nw_{\pm}; n^M x_0) + \delta n^{\frac{1}{5}} \quad \text{for } z \in \Sigma_{\text{curved}} \setminus \left(D_{n^{\frac{3}{5}}}(nw_+) \cup D_{n^{\frac{3}{5}}}(nw_-) \right), \quad (2.19)$$

$$\operatorname{Re} F(z; n^M x_0) \leq \operatorname{Re} F(nw_{\pm}; n^M x_0) - \delta n \left| \frac{z}{n} - w_{\pm} \right|^2 \quad \text{for } z \in \mathcal{C} \cap D_{n^{\frac{3}{5}}}(nw_{\pm}), \quad (2.20)$$

$$\operatorname{Re} F(z; n^M x_0) < \operatorname{Re} F(nw_{\pm}; n^M x_0) - \delta n^{\frac{1}{5}} \quad \text{for } z \in \mathcal{C} \setminus \left(D_{n^{\frac{3}{5}}}(nw_+) \cup D_{n^{\frac{3}{5}}}(nw_-) \right), \quad (2.21)$$

$$\operatorname{Re} F(z; n^M x_0) < \operatorname{Re} F(nw_{\pm}; n^M x_0) - \delta |z| \quad \text{for } z \in \mathcal{C} \cap \{|z| > \delta^{-1}n\}. \quad (2.22)$$

Since the proof of this lemma is lengthy and technical, we decide to postpone it to Section 3.2.

To grasp the meaning of Lemma 2.1 before getting involved in the delicate inequalities, it is better to consult Figure 4, the level line $\operatorname{Re} \hat{F}(z; x_0) = \hat{F}(w_{\pm}; x_0)$. Since the $n \rightarrow \infty$ behaviour of $F(z; n^M x_0)$ is determined by $\hat{F}(z/n; x_0)$, as shown in (2.3), (2.5) and (2.6), by comparing the shapes of Σ_{curved} and \mathcal{C} with the level line in Figure 4, we find that the value $\operatorname{Re} F(z; n^M x_0)$ attains its maximum over \mathcal{C} around nw_{\pm} , while it attains its minimum over Σ_{curved} around these two points.

2.2.2 Evaluation of I_2 as $\epsilon \rightarrow 0$

Note that the contour Σ depends on a parameter ϵ . By taking $\epsilon \rightarrow 0$, we have

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} I_2 &:= \lim_{\epsilon \rightarrow 0} \frac{y^{-1}}{(2\pi i)^2} \int_{\mathcal{C}} ds \int_{\Sigma_3 \cup \Sigma_4} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} \\ &= \frac{y^{-1}}{2\pi i} \int_{nw_-}^{nw_+} \frac{e^{F(s;y)}}{e^{F(s;x)}} ds = \frac{y^{-1}}{2\pi i} \int_{nw_-}^{nw_+} \left(\frac{x}{y} \right)^s ds \\ &= \frac{1}{2\pi i y \log(\frac{x}{y})} \left(\left(\frac{x}{y} \right)^{nw_+} - \left(\frac{x}{y} \right)^{nw_-} \right). \end{aligned} \quad (2.23)$$

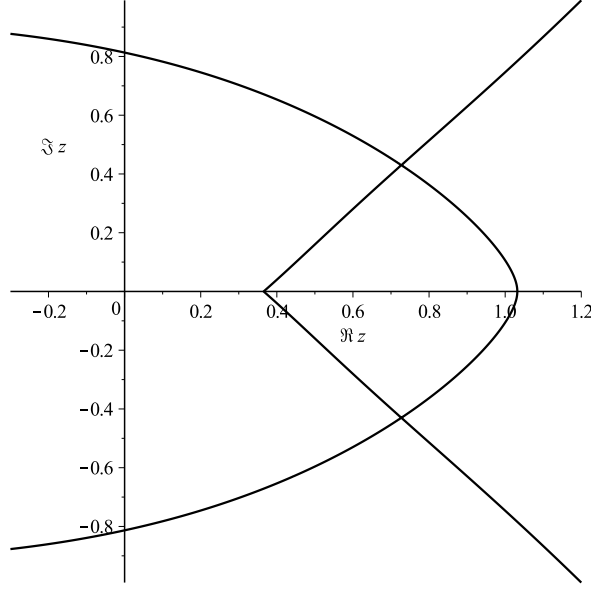


Figure 4: The level line $\text{Re } \hat{F}(z; x_0) = \text{Re } \hat{F}(w_{\pm}; x_0)$ with $M = 3$ and $x_0 = 1$. The level line has two intersections w_{\pm} , and it divides the complex plane into five parts. In the top, central, and bottom parts, $\text{Re } \hat{F}(z; x_0) < \text{Re } \hat{F}(w_{\pm}; x_0)$, while in the left and right parts, $\text{Re } \hat{F}(z; x_0) > \text{Re } \hat{F}(w_{\pm}; x_0)$.

Here we assume that $x \neq y$, and use Cauchy's theorem in the first step. With the values of x, y given in (2.9), it is readily seen that as $n \rightarrow \infty$,

$$\frac{x}{y} = 1 + \frac{\xi - \eta}{n\rho(\varphi)x_0} + \mathcal{O}(n^{-2}), \quad y \log\left(\frac{x}{y}\right) = \frac{n^{M-1}(\xi - \eta)}{\rho(\varphi)} (1 + \mathcal{O}(n^{-1})), \quad (2.24)$$

where φ is related to x_0 by (1.13) and $\rho(\varphi)$ is defined in (1.14). These approximations, together with w_{\pm} given in (2.10), imply that if $\xi \neq \eta$, then

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} I_2 &= \frac{\rho(\varphi)}{2\pi i n^{M-1}(\xi - \eta)(1 + \mathcal{O}(n^{-1}))} \left(e^{\frac{(\xi - \eta)w_+}{\rho(\varphi)x_0}} (1 + \mathcal{O}(n^{-1})) - e^{\frac{(\xi - \eta)w_-}{\rho(\varphi)x_0}} (1 + \mathcal{O}(n^{-1})) \right) \\ &= \frac{e^{\pi\xi \cot \varphi}}{e^{\pi\eta \cot \varphi}} \frac{\rho(\varphi)}{n^{M-1}} \frac{\sin \pi(\xi - \eta)}{\pi(\xi - \eta)} + \mathcal{O}(n^{-M}) \end{aligned} \quad (2.25)$$

for large n . Note that although we define I_2 as a function with real variables x and y , it is also a well defined analytic function if we understand x and y as complex variables. Then (2.25) also holds if ξ and η are distinct complex numbers. By analytic continuation we have that (2.25) also holds for $\xi = \eta$, and particularly for the case that they are identical real numbers.

2.2.3 Evaluation of I_1 as $\epsilon \rightarrow 0$

Parallel to (2.23), by taking the limit $\epsilon \rightarrow 0$, it follows

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} I_1 &:= \lim_{\epsilon \rightarrow 0} \frac{y^{-1}}{(2\pi i)^2} \int_{\mathcal{C}} ds \int_{\Sigma_1 \cup \Sigma_2} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} \\
&= \frac{y^{-1}}{(2\pi i)^2} \lim_{\epsilon \rightarrow 0+} \int_{n\tilde{\Sigma}^r \setminus (D_\epsilon(nw_+) \cup D_\epsilon(nw_-))} \left(\int_{\mathcal{C}} ds \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} \right) dt \\
&= \frac{y^{-1}}{(2\pi i)^2} \text{p.v.} \int_{n\tilde{\Sigma}^r} \left(\int_{\mathcal{C}} ds \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} \right) dt,
\end{aligned} \tag{2.26}$$

where p.v. means the Cauchy principal value. For the definition and properties of Cauchy principal value for contour integral; see [34, Section 8.3, Page 191]. We remark that the integral with respect to t on $n\tilde{\Sigma}^r$ in (2.26) is Riemann integrable, but has discontinuities at nw_{\pm} , the intersections of \mathcal{C} and $n\tilde{\Sigma}^r$. Thus there is no serious integrability problem in the Cauchy principal value for the integral over $n\tilde{\Sigma}^r$.

To estimate the limit of I_1 , we define

$$\mathcal{C}_{\text{local}}^{\pm} = \mathcal{C} \cap D_{n^{\frac{3}{5}}}(nw_{\pm}), \quad \Sigma_{\text{local}}^{\pm} = \Sigma \cap D_{n^{\frac{3}{5}}}(nw_{\pm}), \tag{2.27}$$

and show that the main contribution to the Cauchy principal integral is from $\mathcal{C}_{\text{local}}^+ \times \Sigma_{\text{local}}^+$ and $\mathcal{C}_{\text{local}}^- \times \Sigma_{\text{local}}^-$ in the sense that remaining part of the integral is negligible in the asymptotic analysis.

It is clear that for $s \in \mathcal{C}_{\text{local}}^+$ and $t \in \Sigma_{\text{local}}^+$, we can approximate $F(s; n^M x_0)$ and $F(t; n^M x_0)$ by \tilde{F} as in (2.3) and furthermore by \hat{F} that is defined in (2.6). We make the change of variables

$$s = nw_+ + n^{\frac{1}{2}}u, \quad t = nw_+ + n^{\frac{1}{2}}v. \tag{2.28}$$

It then follows from (2.9), (2.11) and (2.12) that, uniformly for all $s \in D_{n^{3/5}}(nw_+)$,

$$\begin{aligned}
e^{F(s;y)} &= e^{F(s;n^M x_0)} \left(1 + \frac{\eta}{n\rho(\varphi)x_0} \right)^{-s} \\
&= n^n e^{\tilde{c}_M} e^{n\hat{F}(w_+ + n^{-\frac{1}{2}}u; x_0)} \left(1 + \frac{\eta}{n\rho(\varphi)x_0} \right)^{-s} \left(1 + \mathcal{O}(n^{-\frac{1}{2}}) \right) \\
&= n^n e^{\tilde{c}_M + n\hat{F}(w_+; x_0)} e^{\frac{\hat{F}_{zz}(w_+; x_0)}{2} u^2} \left(1 + \frac{\eta}{n\rho(\varphi)x_0} \right)^{-s} \left(1 + \mathcal{O}(n^{-\frac{1}{5}}) \right) \\
&= n^n e^{\tilde{c}_M + n\hat{F}(w_+; x_0)} e^{\frac{\hat{F}_{zz}(w_+; x_0)}{2} u^2} e^{-\frac{w_+ \eta}{x_0 \rho(\varphi)}} \left(1 + \mathcal{O}(n^{-\frac{1}{5}}) \right),
\end{aligned} \tag{2.29}$$

where

$$\tilde{c}_M = \sum_{j=1}^M \left(\nu_j + \frac{1}{2} \right) \log(nw_+) + \frac{1}{2} \log \left(\frac{w_+}{1-w_+} \right) + \frac{M}{2} \log(2\pi). \tag{2.30}$$

A parallel argument yields that uniformly for $t \in D_{n^{3/5}}(nw_+)$,

$$e^{F(t;x)} = n^n e^{\tilde{c}_M + n\hat{F}(w_+; x_0)} e^{\frac{\hat{F}_{zz}(w_+; x_0)}{2} v^2} e^{-\frac{w_+ \xi}{x_0 \rho(\varphi)}} \left(1 + \mathcal{O}(n^{-\frac{1}{5}}) \right). \tag{2.31}$$

As a consequence, (noting that $s - t = \sqrt{n}(u - v)$)

$$\begin{aligned} \text{p.v.} \int_{\mathcal{C}_{\text{local}}^+} ds \oint_{\Sigma_{\text{local}}^+} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} \\ = \frac{e^{-\frac{w_+(\xi-\eta)}{\rho(\varphi)x_0}}}{\sqrt{n}} \text{p.v.} \int_{\mathcal{C}_{\text{local}}^+} ds \oint_{\Sigma_{\text{local}}^+} dt \frac{e^{\frac{\hat{F}_{zz}(w_+;x_0)}{2}u^2}}{e^{\frac{\hat{F}_{zz}(w_+;x_0)}{2}v^2}} \frac{1 + \mathcal{O}(n^{-\frac{1}{5}})}{u-v}, \end{aligned} \quad (2.32)$$

where on the right-hand side, we understand u and v as functions of s and t respectively, as defined by (2.28). Note that the $\mathcal{O}(n^{-1/5})$ term in the integrand on the right-hand side of (2.32) is uniform and analytic in $D_{n^{3/5}}(nw_+)$. Comparing the result of (2.29) with $y = n^M x_0$ and Lemma 2.1, we have that there exists a constant $\epsilon_1 > 0$ such that for all $s \in \mathcal{C}_{\text{local}}^+$,

$$\left| e^{\frac{\hat{F}_{zz}(w_+;x_0)}{2}u^2} \right| \leq e^{-\epsilon_1|u|^2}. \quad (2.33)$$

Similarly, a comparison between (2.31) and (2.18) in Lemma 2.1 implies that there is a constant $\epsilon_2 > 0$ such that for all $t \in \Sigma_{\text{local}}^+$,

$$\left| e^{\frac{\hat{F}_{zz}(w_+;x_0)}{2}v^2} \right| \geq e^{\epsilon_2|v|^2}. \quad (2.34)$$

Hence a standard application of the saddle point method yields

$$\begin{aligned} \text{p.v.} \int_{\mathcal{C}_{\text{local}}^+} ds \oint_{\Sigma_{\text{local}}^+} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} \\ = \lim_{\epsilon \rightarrow 0^+} \int_{\mathcal{C}_{\text{local}}^+} ds \oint_{\Sigma_{\text{local}}^+ \setminus D_\epsilon(nw_+)} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} = \mathcal{O}\left(n^{\frac{1}{2}}\right). \end{aligned} \quad (2.35)$$

In a similar manner, by setting

$$s = nw_- + n^{\frac{1}{2}}u, \quad t = nw_- + n^{\frac{1}{2}}v, \quad (2.36)$$

we have

$$\text{p.v.} \int_{\mathcal{C}_{\text{local}}^-} ds \oint_{\Sigma_{\text{local}}^-} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} = \mathcal{O}\left(n^{\frac{1}{2}}\right). \quad (2.37)$$

Finally, we note by (2.19), (2.21) and (2.22) in Lemma 2.1 that there exists $\epsilon_3 > 0$ such that for large enough n

$$\begin{aligned} \left| e^{-F(t;x)} \right| &= \left| e^{-F(t;n^M x_0)} \left(1 + \frac{\xi}{n\rho(\varphi)x_0} \right)^t \right| < \left| e^{-F(nw_\pm; n^M x_0)} \right| e^{-\epsilon_3 n^{\frac{1}{5}}} \quad \text{if } t \in \Sigma \setminus \Sigma_{\text{local}}^\pm, \\ \left| e^{F(s;y)} \right| &= \left| e^{F(s;n^M x_0)} \left(1 + \frac{\eta}{n\rho(\varphi)x_0} \right)^{-s} \right| < \begin{cases} \left| e^{F(nw_\pm; n^M x_0)} \right| e^{-\epsilon_3 n^{\frac{1}{5}}} & \text{if } s \in \mathcal{C} \setminus \mathcal{C}_{\text{local}}^\pm, \\ \left| e^{F(nw_\pm; n^M x_0)} \right| e^{-\epsilon_3 |s|} & \text{if } s \in \mathcal{C} \cap \{|s| > \frac{n}{\epsilon_3}\}. \end{cases} \end{aligned} \quad (2.38)$$

With the aid of the estimates (2.38), (2.39), (2.29), (2.31), (2.33) and (2.34), we obtain

$$\begin{aligned} \text{p.v.} \int_{\mathcal{C}} ds \oint_{\Sigma} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} - \text{p.v.} \int_{\mathcal{C}_{\text{local}}^+} ds \oint_{\Sigma_{\text{local}}^+} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} \\ - \text{p.v.} \int_{\mathcal{C}_{\text{local}}^-} ds \oint_{\Sigma_{\text{local}}^-} dt \frac{e^{F(s;y)}}{e^{F(t;x)}} \frac{1}{s-t} = \mathcal{O}\left(e^{-\epsilon n^{\frac{1}{5}}}\right). \end{aligned} \quad (2.40)$$

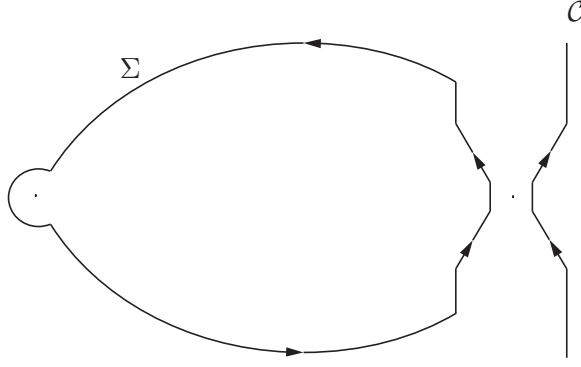


Figure 5: The contours \mathcal{C} and Σ defined in (2.46) and (2.49)

This, together with (2.35), (2.37) and (2.26), implies

$$\lim_{\epsilon \rightarrow 0} I_1 = y^{-1} \left(\mathcal{O}(n^{\frac{1}{2}} + n^{\frac{1}{2}} + \mathcal{O}(e^{-\epsilon n^{\frac{1}{5}}})) \right) = \mathcal{O}(n^{-M+\frac{1}{2}}), \quad (2.41)$$

where we use that $y = \mathcal{O}(n^M)$. Summing up (2.25) and (2.41) and letting $n \rightarrow \infty$, we derive (1.16) and complete the proof of Theorem 1.1.

2.3 Proof of Theorem 1.2

In view of the scalings of x, y in (1.18), we set

$$x = n^M \left(x_* + \frac{c_2 \xi}{n^{2/3}} \right), \quad y = n^M \left(x_* + \frac{c_2 \eta}{n^{2/3}} \right), \quad (2.42)$$

where $\xi, \eta \in \mathbb{R}$,

$$x_* = \frac{(M+1)^{M+1}}{M^M} \quad \text{and} \quad c_2 = \frac{(M+1)^{M+\frac{2}{3}}}{2^{\frac{1}{3}} M^{M-1}},$$

are defined in (1.19). Thus, we write (2.7) as

$$K_n(x, y) = \frac{y^{-1}}{(2\pi i)^2} \int_{\mathcal{C}} ds \oint_{\Sigma} dt \frac{e^{F(s; n^M x_*)} (1 + n^{-\frac{2}{3}} c_1^{-1} \eta)^{-s}}{e^{F(t; n^M x_*)} (1 + n^{-\frac{2}{3}} c_1^{-1} \xi)^{-t}} \frac{1}{s - t} \quad (2.43)$$

with

$$c_1 = \frac{x_*}{c_2} = \frac{2^{\frac{1}{3}} (M+1)^{\frac{1}{3}}}{M}. \quad (2.44)$$

In this case, we will choose the contours \mathcal{C} and Σ in (2.7) such that Σ is on the left hand side of \mathcal{C} , as illustrated in Figure 5. To describe \mathcal{C} , we denote

$$z_0 = 1 + \frac{1}{M}, \quad (2.45)$$

and then define

$$\mathcal{C} = \mathcal{C}_{\text{local}} \cup \mathcal{C}_{\text{global}}, \quad (2.46)$$

where

$$\begin{aligned} \mathcal{C}_{\text{local}} = & \left\{ n z_0 + c_1 n^{\frac{2}{3}} r e^{\pi i/3} \mid r \in \left[1, n^{\frac{1}{30}} \right] \right\} \cup \left\{ n z_0 + c_1 n^{\frac{2}{3}} r e^{-\pi i/3} \mid r \in \left[1, n^{\frac{1}{30}} \right] \right\} \\ & \cup \left\{ n z_0 + \frac{c_1 n^{\frac{2}{3}}}{2} + i c_1 n^{\frac{2}{3}} r \mid r \in \left[-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2} \right] \right\}, \end{aligned} \quad (2.47)$$

and

$$\mathcal{C}_{\text{global}} = \left\{ nz_0 + \frac{1}{2}c_1 n^{\frac{7}{10}} + iy \mid y \in \left(-\infty, -\frac{\sqrt{3}}{2}c_1 n^{\frac{7}{10}} \right] \cup \left[\frac{\sqrt{3}}{2}c_1 n^{\frac{7}{10}}, \infty \right) \right\}. \quad (2.48)$$

The orientation of \mathcal{C} is taken to be upward. The contour Σ is defined as the union of contours

$$\Sigma = \Sigma_{\text{local}} \cup \Sigma_{\text{global}}, \quad \text{and} \quad \Sigma_{\text{global}} = \Sigma_{\text{curved}} \cup \Sigma_{\text{vertical}}. \quad (2.49)$$

The contour Σ_{local} is defined by

$$\begin{aligned} \Sigma_{\text{local}} = & \left\{ nz_0 + c_1 n^{\frac{2}{3}} r e^{2\pi i/3} \mid r \in \left[1, n^{\frac{1}{30}} \right] \right\} \cup \left\{ nz_0 + c_1 n^{\frac{2}{3}} r e^{-2\pi i/3} \mid r \in \left[1, n^{\frac{1}{30}} \right] \right\} \\ & \cup \left\{ nz_0 - \frac{c_1 n^{\frac{2}{3}}}{2} + i c_1 n^{\frac{2}{3}} r \mid r \in \left[-\frac{\sqrt{3}}{2}, \frac{\sqrt{3}}{2} \right] \right\}. \end{aligned} \quad (2.50)$$

The contour Σ_{global} depends on a small constant $\epsilon > 0$. Define

$$r = \frac{[\epsilon n] + \frac{1}{2}}{n}. \quad (2.51)$$

With the contour $\tilde{\Sigma}^r$ constructed in Section 3.1, we denote by $z_{\pm} \in \mathbb{C}_{\pm}$ the two intersection points of $\tilde{\Sigma}^r$ with the vertical line $\text{Re } z = z_0 - \frac{1}{2}c_1 n^{-3/10}$. We then define

$$\Sigma_{\text{curved}} = n\tilde{\Sigma}^r \cap \left\{ \text{Re } z < nz_0 - \frac{1}{2}c_1 n^{\frac{7}{10}} \right\} \quad (2.52)$$

and

$$\Sigma_{\text{vertical}} = \text{two vertical line segments connecting } nz_{\pm} \text{ and } nz_0 \pm c_1 n^{\frac{7}{10}} e^{2\pi i/3}. \quad (2.53)$$

Note that Σ is a closed contour with counterclockwise orientation.

Similar to Lemma 2.1, we have the following properties of $F(z; n^M x_*)$ on the contours \mathcal{C} and Σ .

Lemma 2.2. *There exists a positive constant $\delta > 0$ such that for n large enough,*

$$\text{Re } F(z; n^M x_*) > \text{Re } F(nz_0; n^M x_*) + \delta n^{\frac{1}{10}} \quad \text{for } z \in \Sigma_{\text{global}}, \quad (2.54)$$

$$\text{Re } F(z; n^M x_*) < \text{Re } F(nz_0; n^M x_*) - \delta n^{\frac{1}{10}} \quad \text{for } z \in \mathcal{C}_{\text{global}}, \quad (2.55)$$

$$\text{Re } F(z; n^M x_*) < \text{Re } F(nz_0; n^M x_*) - \delta |z| \quad \text{for } z \in \mathcal{C}_{\text{global}} \cap \{|z| > \delta^{-1} n\}. \quad (2.56)$$

The proof of this lemma is postponed to Section 3.3.

The strategy now is first to consider the double contour integral in (2.43) with \mathcal{C} and Σ restricted to $\mathcal{C}_{\text{local}}$ and Σ_{local} , respectively. It turns out that the integral with the restricted domain yields the Airy kernel in the large n limit. Later we show that the remaining part of the integral is negligible in the asymptotic analysis.

For $s \in \mathcal{C}_{\text{local}}$ and $t \in \Sigma_{\text{local}}$, we can approximate $F(s; n^M x_*)$ and $F(t; n^M x_*)$ by \tilde{F} as in (2.3) and furthermore by \hat{F} that is defined in (2.6). By making the change of variables

$$s = nz_0 + n^{\frac{2}{3}} c_1 u, \quad t = nz_0 + n^{\frac{2}{3}} c_1 v, \quad (2.57)$$

it follows that

$$\begin{aligned} F(s; n^M x_*) &= \tilde{F}(s; n^M x_*) + c_M + \mathcal{O}(n^{-\frac{1}{3}}) \\ &= n\hat{F}(z_0 + n^{-\frac{1}{3}}c_1 u; x_*) + n \log n + c_M + \mathcal{O}(n^{-\frac{1}{3}}), \end{aligned} \quad (2.58)$$

where

$$c_M = \sum_{j=1}^M \left(\nu_j + \frac{1}{2} \right) \log \left(\frac{n(M+1)}{M} \right) + \frac{1}{2} \log(M+1) + \frac{M}{2} \log(2\pi). \quad (2.59)$$

Straightforward calculations show that

$$\hat{F}_z(z_0; x_*) = 0, \quad \hat{F}_{zz}(z_0; x_*) = 0, \quad \hat{F}_{zzz}(z_0; x_*) = \frac{M^3}{M+1}. \quad (2.60)$$

We then obtain from Taylor's expansion of (2.6) that

$$\begin{aligned} &\hat{F}(z_0 + n^{-\frac{1}{3}}c_1 u; x_*) \\ &= \hat{F}(z_0; x_*) + \hat{F}_z(z_0; x_*)c_1 u n^{-\frac{1}{3}} + \frac{1}{2}\hat{F}_{zz}(z_0; x_*)c_1^2 u^2 n^{-\frac{2}{3}} + \frac{1}{6}\hat{F}_{zzz}(z_0; x_*)c_1^3 u^3 n^{-1} + \mathcal{O}(n^{-\frac{6}{5}}) \\ &= \hat{F}(z_0; x_*) + \frac{u^3}{3n} + \mathcal{O}(n^{-\frac{6}{5}}), \end{aligned} \quad (2.61)$$

uniformly valid for $u \in D_{n^{1/30}}(0)$. We also note that, by (2.44), (2.57) and (2.45),

$$\left(1 + n^{-\frac{2}{3}}c_1^{-1}\eta\right)^{-s} = e^{-2^{-\frac{1}{3}}(M+1)^{\frac{2}{3}}\eta n^{\frac{1}{3}}} e^{-u\eta} \left(1 + \mathcal{O}(n^{-\frac{1}{3}})\right), \quad (2.62)$$

for all $u \in D_{n^{1/30}}(0)$ and η in a compact subset of \mathbb{R} . Combining (2.58), (2.61) and (2.62), we find

$$e^{F(s; n^M x_*)} \left(1 + n^{-\frac{2}{3}}c_1^{-1}\eta\right)^{-s} = n^n e^{c_M + \hat{F}(z_0; x_*)n} e^{\frac{1}{3}u^3 - u\eta} e^{-2^{-\frac{1}{3}}(M+1)^{\frac{2}{3}}\eta n^{\frac{1}{3}}} \left(1 + \mathcal{O}(n^{-\frac{1}{5}})\right), \quad (2.63)$$

uniformly for $s \in \mathcal{C}_{\text{local}}$ and η in a compact subset of \mathbb{R} . Similarly, if x and t are expressed respectively by ξ and v via (2.42) and (2.57), where ξ belongs to a compact subset of \mathbb{R} and $t \in \Sigma_{\text{local}}$, we have that uniformly in t and ξ

$$e^{F(t; n^M x_*)} \left(1 + n^{-\frac{2}{3}}c_1^{-1}\xi\right)^{-s} = n^n e^{c_M + \hat{F}(z_0; x_*)n} e^{\frac{1}{3}v^3 - v\xi} e^{-2^{-\frac{1}{3}}(M+1)^{\frac{2}{3}}\xi n^{\frac{1}{3}}} \left(1 + \mathcal{O}(n^{-\frac{1}{5}})\right). \quad (2.64)$$

Substituting (2.63) and (2.64) into the integrand of (2.43), we have

$$\begin{aligned} &\frac{y^{-1}}{(2\pi i)^2} \int_{\mathcal{C}_{\text{local}}} ds \oint_{\Sigma_{\text{local}}} dt \frac{e^{F(s; n^M x_*)} \left(1 + n^{-\frac{2}{3}}c_1^{-1}\eta\right)^{-s}}{e^{F(t; n^M x_*)} \left(1 + n^{-\frac{2}{3}}c_1^{-1}\xi\right)^{-t}} \frac{1}{s-t} \\ &= \frac{e^{2^{-\frac{1}{3}}(M+1)^{\frac{2}{3}}(\xi-\eta)n^{\frac{1}{3}}}}{n^{M-\frac{2}{3}}x_*c_1^{-1}} \left(\frac{1}{(2\pi i)^2} \int_{\mathcal{C}_r} du \int_{\Sigma_r} dv \frac{e^{\frac{1}{3}u^3 - u\eta}}{e^{\frac{1}{3}v^3 - v\xi}} \frac{1}{u-v} + \mathcal{O}(n^{-\frac{1}{5}}) \right) \\ &= \frac{e^{2^{-\frac{1}{3}}(M+1)^{\frac{2}{3}}(\xi-\eta)n^{\frac{1}{3}}}}{n^{M-\frac{2}{3}}c_2} \left(K_{\text{Ai}}(\xi, \eta) + \mathcal{O}(n^{-\frac{1}{5}}) \right), \end{aligned} \quad (2.65)$$

where Σ_r and \mathcal{C}_r are the images of $\mathcal{C}_{\text{local}}$ and Σ_{local} (see (2.47) and (2.50)) under the change of variables (2.57) (see Figure 6 for an illustration), and the last equality follows from the integral representation of Airy kernel shown in (1.17).

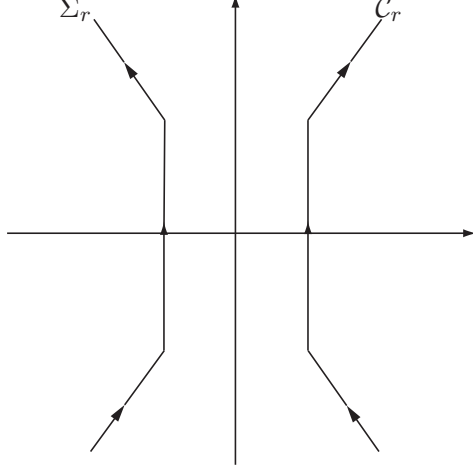


Figure 6: The contours Σ_r and \mathcal{C}_r in (2.65)

In a manner similar to (2.62), we find that

$$e^{2^{-\frac{1}{3}}(M+1)\frac{2}{3}(\eta-\xi)n^{\frac{1}{3}}} \frac{e^{F(s;n^M x_*)} \left(1 + n^{-\frac{2}{3}} c_1^{-1} \eta\right)^{-s}}{e^{F(t;n^M x_*)} \left(1 + n^{-\frac{2}{3}} c_1^{-1} \xi\right)^{-t}} = \frac{e^{F(s;n^M x_*)} \left(1 + n^{-\frac{2}{3}} c_1^{-1} \eta\right)^{-(s-nz_0)}}{e^{F(t;n^M x_*)} \left(1 + n^{-\frac{2}{3}} c_1^{-1} \xi\right)^{-(t-nz_0)}} \left(1 + \mathcal{O}\left(n^{-\frac{1}{3}}\right)\right). \quad (2.66)$$

Then as a consequence of Lemma 2.2, there exists a constant $\delta > 0$ such that for n large enough

$$\left| e^{-F(t;n^M x_*)} \left(1 + n^{-\frac{2}{3}} c_1^{-1} \xi\right)^{(t-nz_0)} \right| < \left| e^{-F(nz_0;n^M x_*)} \right| e^{-\delta n^{\frac{1}{10}}} \quad \text{if } t \in \Sigma_{\text{global}}, \quad (2.67)$$

$$\left| e^{F(s;n^M x_*)} \left(1 + n^{-\frac{2}{3}} c_1^{-1} \eta\right)^{-(s-nz_0)} \right| < \begin{cases} \left| e^{F(nz_0;n^M x_*)} \right| e^{-\delta n^{\frac{1}{10}}} & \text{if } s \in \mathcal{C}_{\text{global}}, \\ \left| e^{F(nz_0;n^M x_*)} \right| e^{-\delta |s|} & \text{if } s \in \mathcal{C}_{\text{global}} \cap \{|s| > \frac{\delta}{n}\}. \end{cases} \quad (2.68)$$

We conclude by (2.66), (2.67), (2.68), and the asymptotics of the integrand of (2.43) given in (2.63) and (2.64) that

$$\frac{y^{-1}}{(2\pi i)^2} \int_{\mathcal{C}} ds \oint_{\Sigma} dt \frac{e^{F(s;x_*)} (1 + n^{-\frac{2}{3}} c_1^{-1} \eta)^{-s}}{e^{F(t;x_*)} (1 + n^{-\frac{2}{3}} c_1^{-1} \xi)^{-t}} \frac{1}{s-t} - \frac{y^{-1}}{(2\pi i)^2} \int_{\mathcal{C}_{\text{local}}} ds \oint_{\Sigma_{\text{local}}} dt \frac{e^{F(s;x_*)} (1 + n^{-\frac{2}{3}} c_1^{-1} \eta)^{-s}}{e^{F(t;x_*)} (1 + n^{-\frac{2}{3}} c_1^{-1} \xi)^{-t}} \frac{1}{s-t} = \mathcal{O}\left(e^{-\delta n^{\frac{1}{10}}}\right). \quad (2.69)$$

A combination of the above formula and (2.65) gives us (1.18), and completes the proof of Theorem 1.2.

3 Contour constructions and proofs of Lemmas 2.1 and 2.2

In this section, we first construct two contours $\tilde{\mathcal{C}}_{x_0}$ and $\tilde{\Sigma}^\epsilon$, from which we can describe precisely the contours of the double integral (2.7) used in the proofs of our main theorems. The

contour $\tilde{\mathcal{C}}_{x_0}$ depends on $x_0 \in (0, x_*)$, where $x_* = (M+1)^{M+1}/M^M$ is defined in (1.19). The other contour $\tilde{\Sigma}^\epsilon$ is dependent on a small parameter $\epsilon > 0$. Two technical lemmas regarding the behavior of the function $\text{Re } \hat{F}$ on these two contours are then proved. With the aid of these two lemmas, we finally finish the proofs of Lemmas 2.1 and 2.2 used in Sections 2.2 and 2.3, respectively.

3.1 Constructions of contours $\tilde{\mathcal{C}}_{x_0}$ and $\tilde{\Sigma}^\epsilon$

Recall that for each $x_0 \in (0, x_*)$, which can be parametrized by $\varphi \in (0, \pi/(M+1))$ as in (1.13), we have two complex conjugate saddle points w_\pm of $\hat{F}(z; x_0)$ defined in (2.10). The contour $\tilde{\mathcal{C}}_{x_0}$ is defined by

$$\tilde{\mathcal{C}}_{x_0} := \{z \in \mathbb{C} \mid \text{Re } z = \text{Re } w_+ = \text{Re } w_-\}, \quad (3.1)$$

i.e., a vertical line passing through $\text{Re } w_\pm$.

For the construction of $\tilde{\Sigma}^\epsilon$, we first define

$$\tilde{\Sigma} := \tilde{\Sigma}_+ \cup \tilde{\Sigma}_-, \quad (3.2)$$

where

$$\tilde{\Sigma}_+ := \left\{ \zeta(\phi) \mid \phi \in \left[0, \frac{\pi}{M+1}\right] \right\}, \quad \tilde{\Sigma}_- := \left\{ \overline{\zeta(\phi)} \mid \phi \in \left[0, \frac{\pi}{M+1}\right] \right\}, \quad (3.3)$$

with

$$\zeta(\phi) = \frac{\sin((M+1)\phi)}{\sin(M\phi)} e^{i\phi}. \quad (3.4)$$

It is easy to check that $\tilde{\Sigma}_\pm$ lies in \mathbb{C}_\pm , passes through w_\pm , and intersects the real line only at 0 when $\phi = \pi/(M+1)$, and at $1 + M^{-1}$ when $\phi = 0$. Furthermore, as ϕ runs from 0 to $\pi/(M+1)$, the value of $|\zeta(\phi)| = \sin((M+1)\phi)/\sin(M\phi)$ decreases, and as $\phi \rightarrow \pi/(M+1)$ from the left,

$$\zeta(\phi) = \frac{\pi - (M+1)\phi}{\sin(\pi M/(M+1))} e^{\frac{i\pi}{M+1}} \left(1 + \mathcal{O}\left(\frac{\pi}{M+1} - \phi\right)\right). \quad (3.5)$$

Thus, for small $\epsilon > 0$, the part of $\tilde{\Sigma}_\pm$ in the disc $D_\epsilon(0)$ is approximated by the line segments $\{z = r e^{\pm i\pi/(M+1)} \mid r \leq \epsilon\}$. A plot of $\tilde{\Sigma}$ is shown in the left picture of Figure 7. Our basic idea is to construct Σ by $n\tilde{\Sigma}$. But the contour $\tilde{\Sigma}$ passes through the origin, which coincides with the poles of integrand in (2.7), we need to make a small deformation of $\tilde{\Sigma}$ around the origin, which gives the following definition of $\tilde{\Sigma}^\epsilon$:

$$\tilde{\Sigma}^\epsilon := \{z \in \tilde{\Sigma} \mid |z| \geq \epsilon\} \cup \text{the arc of } \{|z| = \epsilon\} \text{ connecting } \tilde{\Sigma} \cap \{|z| = \epsilon\} \text{ and through } -\epsilon, \quad (3.6)$$

with counterclockwise orientation. It is clear that $\tilde{\Sigma}^\epsilon$ is a closed contour enclosing the interval $[0, 1]$; see the right picture of Figure 7 for an illustration.

The next two lemmas give the behaviors of $\text{Re } \hat{F}(z; a)$ (defined in (2.6)) on the contours $\tilde{\Sigma}$ and $\tilde{\mathcal{C}}_{x_0}$, which will be essential in our later proofs of Lemmas 2.1 and 2.2.

Lemma 3.1. *For all $x_0 \in (0, x_*)$, which can be parameterized by $\varphi \in (0, \pi/(M+1))$ as in (1.13), there exist constants $\epsilon, \delta > 0$ such that*

$$\text{Re } \hat{F}(z; x_0) \geq \text{Re } \hat{F}(w_\pm; x_0) + \epsilon |z - w_\pm|^2 \quad \text{for } z \in \tilde{\Sigma}^\epsilon \cap D_\delta(w_\pm). \quad (3.7)$$

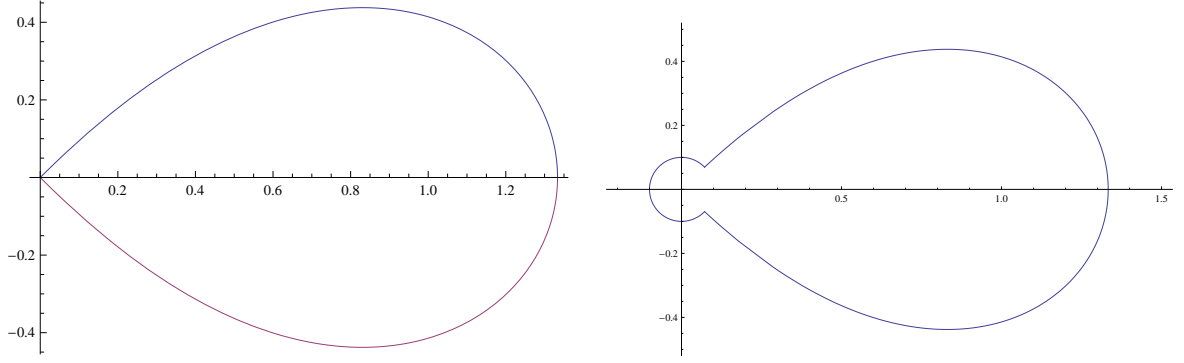


Figure 7: The contours $\tilde{\Sigma}$ (left picture) and $\tilde{\Sigma}^\epsilon$ (right picture) with $M = 3$ and $\epsilon = 0.1$.

Moreover, we have

$$\begin{aligned} \frac{d}{d\phi} \operatorname{Re} \hat{F}(\zeta(\phi); x_0) & \begin{cases} < 0 & \text{for } \phi \in (0, \varphi), \\ > 0 & \text{for } \phi \in \left(\varphi, \frac{\pi}{M+1}\right), \end{cases} \\ \frac{d}{d\phi} \operatorname{Re} \hat{F}(\overline{\zeta(\phi)}; x_0) & \begin{cases} < 0 & \text{for } \phi \in (0, \varphi), \\ > 0 & \text{for } \phi \in \left(\varphi, \frac{\pi}{M+1}\right). \end{cases} \end{aligned} \quad (3.8)$$

We also have

$$\frac{d}{d\phi} \operatorname{Re} \hat{F}(\zeta(\phi); x_*) > 0 \quad \text{and} \quad \frac{d}{d\phi} \operatorname{Re} \hat{F}(\overline{\zeta(\phi)}; x_*) > 0 \quad \text{for } \phi \in \left(0, \frac{\pi}{M+1}\right). \quad (3.9)$$

Proof. Due to the symmetry of $\operatorname{Re} \hat{F}(z; a)$ with respect to the real axis, it suffices to consider the case that $z \in \mathbb{C}_+$, that is, only the inequalities involving in $\zeta(\phi)$.

To show (3.7) and (3.8), we define

$$v(\phi) = \frac{(\sin((M+1)\phi))^{M+1}}{\sin \phi (\sin(M\phi))^M}. \quad (3.10)$$

Note that $x_0 = v(\varphi)$ and for all $\phi \in [0, \pi/(M+1)]$,

$$\zeta(\phi)^{M+1} - v(\phi)(\zeta(\phi) - 1) = 0, \quad (3.11)$$

where $\zeta(\phi)$ is given in (3.4). Thus,

$$\frac{d\hat{F}(\zeta(\phi); x_0)}{d\phi} = \frac{d\hat{F}}{d\zeta} \frac{d\zeta(\phi)}{d\phi} = \log \left(\frac{\zeta(\phi)^{M+1}}{(\zeta(\phi) - 1)x_0} \right) \frac{d\zeta(\phi)}{d\phi} = \log \left(\frac{v(\phi)}{v(\varphi)} \right) \frac{d\zeta(\phi)}{d\phi}. \quad (3.12)$$

Since the function $\sin \theta / \sin(c\theta)$ is a strictly decreasing function on $(0, \pi)$ for $0 < c < 1$, it is readily seen from (3.4) and (3.10) that

$$v(\phi) > 0, \quad \frac{dv(\phi)}{d\phi} < 0, \quad \operatorname{Re} \frac{d\zeta(\phi)}{d\phi} = \frac{d}{d\phi} \left(\frac{\sin((M+1)\phi) \cos \phi}{\sin(M\phi)} \right) < 0, \quad (3.13)$$

for all $\phi \in \left(0, \frac{\pi}{M+1}\right)$. Hence

$$\frac{d\operatorname{Re} \hat{F}(\zeta(\phi); x_0)}{d\phi} = \log \left(\frac{v(\phi)}{v(\varphi)} \right) \frac{d\operatorname{Re} \zeta(\phi)}{d\phi} \begin{cases} < 0 & \text{if } \phi \in (0, \varphi), \\ > 0 & \text{if } \phi \in \left(\varphi, \frac{\pi}{M+1}\right), \end{cases} \quad (3.14)$$

and

$$\left. \frac{d^2 \operatorname{Re} \hat{F}(\zeta(\phi); x_0)}{d\phi^2} \right|_{\phi=\varphi} = \frac{d}{d\phi} \left(\log \left(\frac{v(\phi)}{v(\varphi)} \right) \frac{d \operatorname{Re} \zeta(\phi)}{d\phi} \right) \Big|_{\phi=\varphi} = \frac{d \operatorname{Re} \zeta(\phi)}{d\phi} \Big|_{\phi=\varphi} \frac{v'(\varphi)}{v(\varphi)} > 0, \quad (3.15)$$

which gives us (3.7) and (3.8) for z (or $\zeta(\phi)$) $\in \mathbb{C}_+$.

Finally, note that the inequality (3.9) is the limiting case of (3.8) as $x_0 \rightarrow x_*$, or equivalently, $\varphi \rightarrow 0$, the result is then immediate. \square

Lemma 3.2. *For all conjugate pairs $w_{\pm} \in \mathbb{C}_{\pm}$ locating on $\tilde{\Sigma}$, there exists constants $\epsilon, \delta > 0$ such that for all $a \in \mathbb{R}$*

$$\operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) \leq \operatorname{Re} \hat{F}(w_{\pm}; a) - \epsilon |y - \operatorname{Im} w_{\pm}|^2 \quad \text{for } |y - \operatorname{Im} w_{\pm}| \leq \delta. \quad (3.16)$$

Moreover, we have

$$\begin{aligned} \frac{d}{dy} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) & \begin{cases} < 0 & \text{if } y > \operatorname{Im} w_+, \\ > 0 & \text{if } y \in (0, \operatorname{Im} w_+), \\ < 0 & \text{if } y \in (\operatorname{Im} w_-, 0), \\ > 0 & \text{if } y < \operatorname{Im} w_-, \end{cases} \\ \lim_{y \rightarrow \pm\infty} \frac{d}{dy} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) & = \mp\infty. \end{aligned} \quad (3.17)$$

We also have, for all $c > 0$,

$$\begin{aligned} \frac{d}{dy} \operatorname{Re} \hat{F}(1 + M^{-1} + c + iy; a) & \begin{cases} < 0 & \text{if } y > 0, \\ > 0 & \text{if } y < 0, \end{cases} \\ \lim_{y \rightarrow \pm\infty} \frac{d}{dy} \operatorname{Re} \hat{F}(1 + M^{-1} + c + iy; a) & = \mp\infty. \end{aligned} \quad (3.18)$$

Proof. Similar to the proof of Lemma 3.1, we need only to prove (3.16)–(3.18) for $y > 0$.

By Cauchy-Riemann equations, it follows that, for any $x \in \mathbb{R}$ and $y \in \mathbb{R}_+$,

$$\begin{aligned} \frac{\partial}{\partial y} \operatorname{Re} \hat{F}(x + iy; a) & = -\operatorname{Im} \frac{d}{dz} \hat{F}(z; a) \Big|_{z=x+iy}, \\ \frac{\partial^2}{\partial y^2} \operatorname{Re} \hat{F}(x + iy; a) & = -\operatorname{Re} \frac{d^2}{dz^2} \hat{F}(z; a) \Big|_{z=x+iy}. \end{aligned} \quad (3.19)$$

Since

$$\hat{F}''(z; a) = \frac{M+1}{z} - \frac{1}{z-1}, \quad (3.20)$$

we have

$$\frac{\partial^2}{\partial y^2} \operatorname{Re} \hat{F}(x + iy; a) = \frac{(M+1-Mx)x(x-1) - (Mx+1)y^2}{(x^2+y^2)((x-1)^2+y^2)}, \quad (3.21)$$

which is independent of the parameter a .

To show (3.16) and (3.17), we observe from (2.10) that

$$\operatorname{Re} w_{\pm} = \frac{\sin((M+1)\varphi) \cos \varphi}{\sin(M\varphi)} \in \left(0, \frac{M+1}{M}\right), \quad \varphi \in \left(0, \frac{\pi}{M+1}\right). \quad (3.22)$$

In the case that $0 < \operatorname{Re} w_{\pm} \leq 1$, we have $M + 1 - M\operatorname{Re} w_{\pm} > 0$. This, together with (3.21), implies that

$$\frac{\partial^2}{\partial y^2} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) < 0 \quad \text{for all } y > 0 \quad (3.23)$$

and

$$\lim_{y \rightarrow +\infty} \frac{\partial^2}{\partial y^2} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) = 0. \quad (3.24)$$

Furthermore, since the parameter a is assumed to be real, the value of $\frac{\partial}{\partial y} \operatorname{Re} \hat{F}(x + iy; a)$ does not depend on a . By (3.19) and (2.11), we have

$$\left. \frac{\partial}{\partial y} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) \right|_{y=\operatorname{Im} w_{+}} = -\operatorname{Im} \left. \frac{d}{dz} \hat{F}(z; a) \right|_{z=w_{+}} = -\operatorname{Im} \left. \frac{d}{dz} \hat{F}(z; x_0) \right|_{z=w_{+}} = 0, \quad (3.25)$$

for any $a \in \mathbb{R}$. Thus, $\operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + y; a)$, as a function of $y > 0$, has a critical point at $\operatorname{Im} w_{+}$, and by (3.23), is a concave function attaining its maximum at $\operatorname{Im} w_{+}$. We thus prove (3.16) and (3.17) in this case.

In the case that $1 < \operatorname{Re} w_{\pm} < (M + 1)/M$, the equation $\frac{\partial^2}{\partial y^2} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) = 0$ has a unique real root at

$$y^* = \sqrt{\frac{(M + 1 - M\operatorname{Re} w_{\pm})\operatorname{Re} w_{\pm}(\operatorname{Re} w_{\pm} - 1)}{M\operatorname{Re} w_{\pm} + 1}}, \quad (3.26)$$

for $y \in [0, \infty)$. Thus, $\frac{\partial}{\partial y} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a)$ is strictly increasing if $y \in [0, y^*)$, and strictly decreasing if $y \in (y^*, \infty)$. Note that $\frac{\partial}{\partial y} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a)$ is a continuous odd function in y , one has $\left. \frac{\partial}{\partial y} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) \right|_{y=0} = 0$. Therefore,

$$\frac{\partial}{\partial y} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a) > 0, \quad y \in (0, y^*). \quad (3.27)$$

On the other hand, by (3.25), we have that $\frac{\partial}{\partial y} \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a)$ vanishes at $\operatorname{Im} w_{+}$. Thus we conclude that $\operatorname{Im} w_{+} \in (y^*, \infty)$, and have that on the interval $[y^*, \infty)$, the function $\operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm} + iy; a)$ has a critical point at $\operatorname{Im} w_{+}$, and is a concave function with the maximum at $\operatorname{Im} w_{+}$. Note that (3.24) also holds in this case. We thus prove (3.16) and (3.17) in this case.

We finally prove (3.18). By substituting $x = 1 + M^{-1} + c$ into (3.21), we have

$$\frac{\partial^2}{\partial y^2} \operatorname{Re} \hat{F}(1 + M^{-1} + c + iy; a) < 0 \quad \text{for all } y > 0. \quad (3.28)$$

On the other hand, since $\frac{\partial}{\partial y} \operatorname{Re} \hat{F}(1 + M^{-1} + c + iy; a)$ is a continuous odd function in y , its value at 0 is 0. We conclude that $\frac{\partial}{\partial y} \operatorname{Re} \hat{F}(1 + M^{-1} + c + iy; a) < 0$ for all $y > 0$, and this gives us (3.18). \square

In Sections 2.2 and 2.3, the contours \mathcal{C} and Σ in (2.7) are constructed from $\tilde{\mathcal{C}}_{x_0}$ and $\tilde{\Sigma}^r$, where r depends on n and a small parameter. In the proofs of our main theorems, we need to estimate some integrals over specified contours, which relies on Lemmas 2.1 and 2.2 concerning the inequalities satisfied by $\operatorname{Re} F$ over \mathcal{C} and Σ . We are now ready to prove these two lemmas based on Lemmas 3.1 and 3.2.

3.2 Proof of Lemma 2.1

For notational convenience, we shall write $F(z; n^M x_0)$ as $F(z)$ throughout this subsection.

Proof of (2.18) and (2.19) Recall the contour Σ_{curved} defined by (2.15) and (2.16), we further write it as

$$\Sigma_{\text{curved}} = \Sigma_{\text{left}} \cup \Sigma_{\text{right}}, \quad (3.29)$$

where

$$\Sigma_{\text{left}} := \{z \in \Sigma_{\text{curved}} \mid |z| = nr\}, \quad \Sigma_{\text{right}} := \{z \in \Sigma_{\text{curved}} \mid |z| > nr\}, \quad (3.30)$$

i.e., Σ_{left} is part of a circle centring at 0 with radius nr , and Σ_{right} is the part of Σ_{curved} that does not overlap the circle.

If $z \in \Sigma_{\text{right}}$, it can be expressed as $z = n\zeta(\phi)$ or $z = n\overline{\zeta(\phi)}$ for some $\phi \in (0, \pi/(M+1))$ by (3.4), so there exists a constant $\varepsilon' > 0$, such that for large enough n , $\arg(z + \nu_j + 1) \in (-\pi + \varepsilon', \pi - \varepsilon')$ and $\arg(z - n + 1) \in (-\pi + \varepsilon', \pi - \varepsilon')$. We then apply the Stirling's formula (2.2) to $\Gamma(z + \nu_j + 1)$ and $\Gamma(z - n + 1)$ in formula (2.1), and obtain a uniform approximation of $F(z)$ by $n\hat{F}(z/n; x_0)$, on account of (2.2)–(2.6). Thus, the inequalities (3.7) and (3.8) for $\text{Re } \hat{F}(z; x_0)$ on $\tilde{\Sigma}$ yield the desired inequalities (2.18) and (2.19) for $z \in \Sigma_{\text{right}}$.

If $z \in \Sigma_{\text{left}}$, Stirling's formula (2.2) may not be valid anymore, and we need to pay special attention. Note that there exists a constant $\varepsilon' > 0$, such that for all n large enough, $\arg(-z - \nu_j) \in (-\pi + \varepsilon', \pi - \varepsilon')$ and $\arg(n - z) \in (-\pi + \varepsilon', \pi - \varepsilon')$. We make use of the reflection formula of gamma function

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)} \quad (3.31)$$

to obtain a uniform approximation of $F(z)$. Since

$$\sin(\pi(z - n + 1)) = \pm \sin(\pi z), \quad \sin(\pi(z + \nu_j + 1)) = \pm \sin(\pi z), \quad (3.32)$$

we have

$$\begin{aligned} \text{Re } F(z) &= \text{Re } \log \left(\frac{\Gamma(n-z)}{\prod_{j=0}^M \Gamma(-z - \nu_j)} \frac{\pi^M \sin(\pi(z - n + 1))}{\prod_{j=0}^M \sin(\pi(z + \nu_j + 1))} (n^M x_0)^{-z} \right) \\ &= \text{Re } \log \left(\frac{\Gamma(n-z)}{\prod_{j=0}^M \Gamma(-z - \nu_j)} (n^M x_0)^{-z} \right) - M \log \frac{|\sin(\pi z)|}{\pi} \\ &= \text{Re } \tilde{G}(z; n^M x_0) - M \log |2 \sin(\pi z)| \\ &\quad + \sum_{j=0}^M \left(\nu_j + \frac{1}{2} \right) \log |z| - \frac{1}{2} \log |z - n| + \frac{M}{2} \log(2\pi) + \mathcal{O}(n^{-1}), \end{aligned} \quad (3.33)$$

where

$$\tilde{G}(z; n^M x_0) = (M+1)z(\log(-z) - 1) - (z-n)(\log(n-z) - 1) - (M \log n + \log x_0)z. \quad (3.34)$$

It is also straightforward to check that

$$\tilde{G}(z; n^M x_0) = n\hat{G}\left(\frac{z}{n}; x_0\right) + n \log n, \quad (3.35)$$

where

$$\hat{G}(\zeta; x_0) = (M+1)\zeta(\log(-\zeta) - 1) - (\zeta-1)(\log(1-\zeta) - 1) - \zeta \log x_0. \quad (3.36)$$

Formulas (3.33)–(3.36) constitute a uniform approximation of $\text{Re } F(z)$ for $z \in \Sigma_{\text{left}}$. Now we choose the parameter ε' in (2.16) small enough such that $nw_{\pm} \in \Sigma_{\text{right}}$, thus $\text{Re } F(nw_{\pm})$ can be approximated by (2.3)–(2.6). The inequalities (2.18) and (2.19) follow if we can show that there exists a constant $c > 0$ such that for all large enough n and $z \in \Sigma_{\text{left}}$,

$$\text{Re } \hat{G}\left(\frac{z}{n}; x_0\right) - \frac{M}{n} \log |2 \sin(\pi z)| > \text{Re } \hat{F}(w_{\pm}; x_0) + c. \quad (3.37)$$

To prove (3.37), we note that, by (3.8) in Lemma 3.1,

$$\operatorname{Re} \hat{G}(0; x_0) = \operatorname{Re} \hat{F}(0; x_0) > \operatorname{Re} \hat{F}(w_{\pm}; x_0), \quad (3.38)$$

so we simply take

$$c = \frac{1}{3} \left(\operatorname{Re} \hat{F}(w_{\pm}; x_0) - \operatorname{Re} \hat{F}(0; x_0) \right). \quad (3.39)$$

Since $\operatorname{Re} \hat{G}(\zeta; x_0)$ is a continuous function in the vicinity of 0, we have that if ϵ is small enough, or equivalently, r is small enough, $|\operatorname{Re} \hat{G}(z/n; x_0) - \operatorname{Re} \hat{G}(0; x_0)| < c$ for all $z \in \Sigma_{\text{left}}$. On the other hand, it is straightforward to check that if ϵ' is small enough, and n is large enough, then $Mn^{-1} \log|2 \sin(\pi z)| < c$ for all $z \in \Sigma_{\text{left}}$. Thus (3.37) holds if ϵ' is small enough while n is large enough. This completes the proof of (2.18) and (2.19).

Proof of (2.20)–(2.22) For any $x_0 \in (0, (M+1)^{M+1}/M^M)$, the associated complex conjugate numbers w_{\pm} satisfying $\operatorname{Re} w_{\pm} \in (0, 1 + M^{-1})$; see (3.22). We prove the inequalities in three cases depending on the value of $\operatorname{Re} w_{\pm}$.

We first consider the case that $\operatorname{Re} w_{\pm} > 1$, or equivalently, the vertical contour \mathcal{C} defined in (2.13) is on the right of n . Then for all $z \in \mathcal{C}$, there exists a constant $\epsilon' > 0$ such that $\arg(z + \nu_j + 1) \in (-\pi + \epsilon', \pi - \epsilon')$ and $\arg(z - n + 1) \in (-\pi + \epsilon', \pi - \epsilon')$ for large n enough. The formulas (2.3)–(2.6) then give a uniform approximation of $F(z)$ by $n\hat{F}(z/n; x_0)$, similar to the case that $z \in \Sigma_{\text{right}}$ discussed previously. Hence, (2.20)–(2.22) are direct consequence of (3.16) and (3.17) in Lemma 3.2.

In the case that $\operatorname{Re} w_{\pm} \in (0, 1)$, or equivalently, the vertical contour \mathcal{C} lies between 0 and n , we divide

$$\mathcal{C} = \mathcal{C}_{\text{outer}} \cup \mathcal{C}_{\text{inner}}, \quad (3.40)$$

where

$$\mathcal{C}_{\text{outer}} = \{z \in \mathcal{C} \mid |\operatorname{Im} z| > n\epsilon'\}, \quad \mathcal{C}_{\text{inner}} = \{z \in \mathcal{C} \mid |\operatorname{Im} z| \leq n\epsilon'\} \quad (3.41)$$

and ϵ' is a small positive number.

For $z \in \mathcal{C}_{\text{outer}}$, we can still use the Stirling's formula directly and approximate $F(z)$ by $n\hat{F}(z/n; x_0)$ through (2.3)–(2.6). The desired inequalities (2.20)–(2.22) for such z again follow from (3.16) and (3.17) in Lemma 3.2.

For $z \in \mathcal{C}_{\text{inner}}$, we encounter the problem of validity of Stirling's formula for $\Gamma(z - n + 1)$. With the aid of the reflection formula (3.31), for n large enough, we obtain the following uniform approximation of $\operatorname{Re} F(z)$ given by

$$\begin{aligned} \operatorname{Re} F(z) &= \operatorname{Re} \tilde{H}(z; n^M x_0) + \log|2 \sin(\pi z)| \\ &\quad + \sum_{j=0}^M \left(\nu_j + \frac{1}{2} \right) \log |z| - \frac{1}{2} \log |z - n| + \frac{M}{2} \log(2\pi) + \mathcal{O}(n^{-1}), \end{aligned} \quad (3.42)$$

where

$$\begin{aligned} \tilde{H}(z; n^M x_0) &= (M+1)z(\log z - 1) - (z - n)(\log(n - z) - 1) - (M \log n + \log x_0)z \\ &= n\hat{H}\left(\frac{z}{n}; x_0\right) + n \log n, \end{aligned} \quad (3.43)$$

and

$$\hat{H}(\zeta; x_0) = (M+1)\zeta(\log \zeta - 1) - (\zeta - 1)(\log(1 - \zeta) - 1) - \zeta \log x_0. \quad (3.44)$$

Similar to the discussions used in the proof of (2.18) and (2.19) with $z \in \Sigma_{\text{left}}$, we only need to show that for $z \in \mathcal{C}_{\text{inner}}$, there exists a constant $c > 0$ such that

$$\operatorname{Re} \hat{H}\left(\frac{z}{n}; x_0\right) + \frac{1}{n} \log|2 \sin(\pi z)| < \operatorname{Re} \hat{F}(w_{\pm}; x_0) - c. \quad (3.45)$$

Now we take

$$c = \frac{1}{3} \left(\operatorname{Re} \hat{F}(w_{\pm}; x_0) - \operatorname{Re} \hat{F}(\operatorname{Re} w_{\pm}; x_0) \right), \quad (3.46)$$

which is positive by (3.17) in Lemma 3.1. Since $\operatorname{Re} \hat{H}(\zeta; x_0)$ is continuous in the vicinity of $\operatorname{Re} w_{\pm}$, we have that $|\operatorname{Re} \hat{H}(z/n; x_0) - \operatorname{Re} \hat{H}(\operatorname{Re} w_{\pm}; x_0)| < c$ for all $z \in \mathcal{C}_{\text{inner}}$ if ϵ' is small enough. On the other hand, it is straightforward to check that if ϵ' is small enough and n large enough, then $n^{-1} \log|2 \sin(\pi z)| < c$ for all $z \in \mathcal{C}_{\text{inner}}$. This gives us (2.21) for $z \in \mathcal{C}_{\text{inner}}$, and finishes the proof in this case.

Finally, if $\operatorname{Re} w_{\pm} = 1$, we still divide \mathcal{C} into $\mathcal{C}_{\text{outer}}$ and $\mathcal{C}_{\text{inner}}$ as in (3.40). The estimate of $\operatorname{Re} F(z)$ on $\mathcal{C}_{\text{outer}}$ can be derived from the Stirling's formula, but for $z \in \mathcal{C}_{\text{inner}}$, we need to control the value of $\Gamma(z - n + 1)$ for $z - n = o(n)$. Since the strategy is similar, we omit the details here.

3.3 Proof of Lemma 2.2

For notational convenience, we shall write $F(z; n^M x_*)$ as $F(z)$ throughout this subsection.

Proof of (2.54) For $z \in \Sigma_{\text{curved}}$, the proof is parallel to that of (2.19). The only difference is that after approximating $F(z)$ uniformly by $n\hat{F}(z/n; x_*)$ (defined in (2.6)) or by $n\hat{G}(z/n; x_*)$ (defined in (3.36), with x_0 replaced by x_*), depending on whether $|z| > nr$ or $|z| = nr$, we compare $\hat{F}(z/n; x_*)$ and $\hat{G}(z/n; x_*)$ with $\operatorname{Re} \hat{F}(1 + M^{-1}; x_*)$, instead of $\operatorname{Re} \hat{F}(w_{\pm}; x_*)$ used in the proof of (2.19). We then apply the inequality (3.9), instead of the inequality (3.8), in the comparison. The details are left to the interested readers.

For $z \in \Sigma_{\text{vertical}}$, we apply the approximation of $F(z)$ by $n\hat{F}(z/n; x_*)$ as in (2.2)–(2.6), and reduce the proof of (2.54) for $z \in \Sigma_{\text{vertical}}$ to proving

$$\operatorname{Re} \hat{F}\left(\frac{z}{n}; x_*\right) > \operatorname{Re} \hat{F}(1 + M^{-1}; x_*) + \delta n^{-\frac{9}{10}}, \quad \delta > 0. \quad (3.47)$$

Without loss of generality, we show (3.47) only for $z \in \Sigma_{\text{vertical}} \cap \mathbb{C}_+$. By (3.17) in Lemma 3.2, $\operatorname{Re} \hat{F}(z/n; x_*)$ increases as $\operatorname{Im} z$ increases for $z \in \Sigma_{\text{vertical}} \cap \mathbb{C}_+$. So we only need to check that (3.47) holds for $z = nz_0 + c_1 n^{\frac{7}{10}} e^{2\pi i/3}$, i.e., the lowest end of $\Sigma_{\text{vertical}} \cap \mathbb{C}_+$. The explicit computation in (2.61) gives the approximation of $\operatorname{Re} \hat{F}(z/n; x_*)$ at this point and finishes the proof in this case.

Proof of (2.55) and (2.56) For all $z \in \mathcal{C}_{\text{global}}$, the uniform approximation of $F(z)$ by $n\hat{F}(z/n; x_*)$ as in (2.2)–(2.6) is valid. Then we reduce (2.55) and (2.56) to

$$\operatorname{Re} \hat{F}(z/n; x_*) < \operatorname{Re} \hat{F}(z_0; x_*) - \delta n^{-\frac{9}{10}} \quad \text{for } z \in \mathcal{C}_{\text{global}}, \quad (3.48)$$

$$\operatorname{Re} \hat{F}(z/n; x_*) < \operatorname{Re} \hat{F}(z_0; x_*) - \delta |z|/n \quad \text{for } z \in \mathcal{C}_{\text{global}} \text{ and } |z| > \delta^{-1}n. \quad (3.49)$$

The inequality (3.49) is a direct consequence of (3.18). To prove (3.48) for $z \in \mathcal{C}_{\text{global}} \cap \mathbb{C}_+$, we note that $\operatorname{Re} \hat{F}(z/n; x_*)$ decreases as $\operatorname{Im} z$ increases, as shown in (3.18). Thus we only need to check (3.48) at $z = nz_0 + c_1 n^{\frac{7}{10}} e^{\pi i/3}$, the lowest end of $\mathcal{C}_{\text{global}} \cap \mathbb{C}_+$. The explicit computation (2.61) gives the approximation of $\operatorname{Re} \hat{F}(z/n; x_*)$ at this point and finishes the proof in this case. The inequality (3.48) for $z \in \mathcal{C}_{\text{global}} \cap \mathbb{C}_-$ can be proved in the same way.

4 Bulk and soft edge universality in other product models

As mentioned in Section 1.5, our strategy presented before is not restricted to the particular model, but applicable to other interesting models of products of random matrices. In this section, we demonstrate this aspect by establishing bulk and soft edge universality in two more examples with sketched proofs. One example is the product of M Ginibre matrices and the inverse of K Ginibre matrices studied by Forrester [22], and the other example is the product of $M-1$ Ginibre matrices with one truncated unitary matrix considered by Kuijlaars and Stivigny [37]. Our method can be applied to more cases, notably the newly analysed model in [24] and models that can be expressed in the general double contour integral formalism in [19].

In this section, we use the same notations as in previous sections for objects in different models that have counterpart in the model introduced and computed in Sections 1–3. We hope these notations show the readers analogue in our arguments while do not bring confusion.

4.1 Products of Ginibre matrices and their inverses

This model refers to the product

$$Y_{M,K} = X_M \cdots X_1 (\tilde{X}_K \cdots \tilde{X}_1)^{-1}, \quad (4.1)$$

where X_j , $j = 1, \dots, M$, and \tilde{X}_k , $k = 1, \dots, K$, are complex Ginibre random matrices with size $(n + \nu_j) \times (n + \nu_{j-1})$ and $(n + \tilde{\nu}_k) \times (n + \tilde{\nu}_{k-1})$, respectively. We assume that

$$\nu_0 = \tilde{\nu}_0 = \tilde{\nu}_K = 0, \quad \nu_j, \tilde{\nu}_k \geq 0, \quad (4.2)$$

thus, $Y_{M,K}$ is a rectangular matrix of size $(n + \nu_M) \times n$. Clearly, $Y_{M,K}$ extends products of Ginibre matrices Y_M defined in (1.1).

It was shown in [22, Proposition 5] that the squared singular values of $Y_{M,K}$ forms a determinantal process with the correlation kernel

$$K_n(x, y) = \frac{1}{(2\pi i)^2} \int_{-1/2-i\infty}^{-1/2+i\infty} ds \oint_{\Sigma} dt \prod_{j=0}^M \frac{\Gamma(s + \nu_j + 1)}{\Gamma(t + \nu_j + 1)} \prod_{k=1}^K \frac{\Gamma(n - s + \tilde{\nu}_k)}{\Gamma(n - t + \tilde{\nu}_k)} \frac{\Gamma(t - n + 1)}{\Gamma(s - n + 1)} \frac{x^t y^{-s-1}}{s - t}, \quad (4.3)$$

where Σ is a closed contour going around $0, 1, \dots, n-1$ in the positive direction and $\operatorname{Re} t > -1/2$ for $t \in \Sigma$.

The special case $(M, K) = (M, 0)$ is the model we considered in Sections 1–3, while the special case $(M, K) = (0, K)$ is equivalent spectrally to the model $(K, 0)$ by reciprocal transform. In the generic case $K > 0, M > 0$, the limiting mean density is supported over the whole positive real axis as $n \rightarrow \infty$ (see [22, 23]), which implies that no soft edge occurs. Below we only consider the $K > 0, M > 0$ case.

To state our result for bulk universality, we need the following parametrization of the spectral parameter x_0

$$x_0 = \frac{\left(\sin \left(\frac{M+1}{K+1} \varphi + \frac{K}{K+1} \pi \right) \right)^{M+1}}{(\sin \varphi)^{K+1} \left(\sin \left(\frac{M-K}{K+1} \varphi + \frac{K}{K+1} \pi \right) \right)^{M-K}}, \quad 0 < \varphi < \frac{\pi}{M+1}, \quad (4.4)$$

which is a one-to-one mapping from $(0, \pi/(M+1))$ to $(0, +\infty)$; see [23, 29].

Theorem 4.1 (Bulk universality). *Let $K_n(x, y)$ be the correlation kernel defined in (4.3). For $x_0 \in (0, +\infty)$, which is parametrized by $\varphi \in (0, \pi/(M+1))$ through (4.4), we have, with $\nu_j, \tilde{\nu}_k$ being fixed,*

$$\lim_{n \rightarrow \infty} \frac{e^{-\pi\xi \cot \varphi}}{e^{-\pi\eta \cot \varphi}} \frac{n^{M-K-1}}{\rho(\varphi)} K_n \left(n^{M-K} \left(x_0 + \frac{\xi}{n\rho(\varphi)} \right), n^{M-K} \left(x_0 + \frac{\eta}{n\rho(\varphi)} \right) \right) = K_{\sin}(\xi, \eta) \quad (4.5)$$

uniformly for ξ and η in any compact subset of \mathbb{R} , where the function ρ is given by

$$\rho(\varphi) = \frac{1}{\pi x_0} \frac{\sin \left(\frac{M+1}{K+1} \varphi + \frac{K}{K+1} \pi \right)}{\sin \left(\frac{M-K}{K+1} \varphi + \frac{K}{K+1} \pi \right)} \sin \varphi. \quad (4.6)$$

We now give a sketched proof of the above theorem with emphasis on the key steps.

Sketched proof of Theorem 4.1. We scale the values of x and y in (4.3) such that

$$x = n^{M-K} \left(x_0 + \frac{\xi}{n\rho(\varphi)} \right), \quad y = n^{M-K} \left(x_0 + \frac{\eta}{n\rho(\varphi)} \right), \quad (4.7)$$

where $\xi, \eta \in \mathbb{R}$ and $\rho(\varphi)$ is given in (4.6). By Stirling's formula (2.2) and the reflection formula, it follows that, for n large,

$$K_n(x, y) \sim -\frac{n^{-M+K}}{(2\pi i)^2} \int_{\mathcal{C}} ds \oint_{\Sigma} dt \frac{e^{n(\hat{F}(ns; x_0) - \hat{F}(nt; x_0))}}{s-t} \left(1 + \frac{\xi}{nx_0\rho} \right)^t \left(1 + \frac{\eta}{nx_0\rho} \right)^{-s} \left(x_0 + \frac{\eta}{\rho n} \right)^{-1} \exp \left\{ \sum_{j=0}^M \left(\nu_j + \frac{1}{2} \right) \log \frac{s}{t} + \sum_{k=1}^K \left(\tilde{\nu}_k - \frac{1}{2} \right) \log \frac{n-s}{n-t} - \frac{1}{2} \log \frac{s-n}{t-n} \right\}, \quad (4.8)$$

where the shapes of the contours \mathcal{C} and Σ are to be described later. Here,

$$\hat{F}(z; x_0) = (M+1)(z \log z - 1) + K(1-z)(\log(1-z) - 1) - (z-1)(\log(z-1) - 1) - z \log x_0. \quad (4.9)$$

Since

$$\hat{F}_z(z; x_0) = (M+1) \log z - K \log(1-z) - \log(z-1) - \log x_0, \quad (4.10)$$

the saddle point of $\hat{F}(z; x_0)$ satisfies the following algebraic equation

$$z^{M+1} + x_0(1-z)^{K+1} = 0. \quad (4.11)$$

Particularly, with the help of parametrization (4.4), two solutions of (4.11) can be given explicitly by

$$w_{\pm} = \frac{\sin \left(\frac{M+1}{K+1} \varphi + \frac{K}{K+1} \pi \right)}{\sin \left(\frac{M-K}{K+1} \varphi + \frac{K}{K+1} \pi \right)} e^{\pm i\varphi}; \quad (4.12)$$

see [23] for more details. These two complex conjugate numbers play the same role of w_{\pm} used in Section 2.2.

Similar to the contours used in Section 2.2, the contour \mathcal{C} is chosen to be the straight line

$$\mathcal{C} = \left\{ z \mid \operatorname{Re} \frac{z}{n} = \operatorname{Re} w_{\pm} = \frac{\sin \left(\frac{M+1}{K+1} \varphi + \frac{K}{K+1} \pi \right)}{\sin \left(\frac{M-K}{K+1} \varphi + \frac{K}{K+1} \pi \right)} \cos \varphi \right\}, \quad (4.13)$$

while Σ is chosen to be a deformation based on the contour

$$\tilde{\Sigma} = \left\{ z = \frac{\sin\left(\frac{M+1}{K+1}\phi + \frac{K}{K+1}\pi\right)}{\sin\left(\frac{M-K}{K+1}\phi + \frac{K}{K+1}\pi\right)} e^{i\phi} \mid -\frac{\pi}{M+1} \leq \phi \leq \frac{\pi}{M+1} \right\}, \quad (4.14)$$

in the same way as the Σ in Section 2.2.

One can then show that, in a manner similar to Lemma 2.1, $\text{Re } \hat{F}(z; x_0)$ defined in (4.9) attains its global maximum at $z = w_{\pm}$ for $nz \in \mathcal{C}$ and its global minimum at $z = w_{\pm}$ for $z \in \tilde{\Sigma}$. This in turn implies that the main contribution of the integral in (4.3) comes from the counterpart of the integral I_2 in (2.23), that is,

$$\begin{aligned} K_n(x, y) &\sim \frac{n^{-M+K}}{2\pi i x_0} \int_{nw_-}^{nw_+} ds \left(1 + \frac{\xi}{x_0 \rho n}\right)^s \left(1 + \frac{\eta}{x_0 \rho n}\right)^{-s} \\ &\sim n^{-M+K+1} \rho e^{\pi(\xi-\eta) \cot \varphi} \frac{\sin \pi(\xi - \eta)}{\pi(\xi - \eta)}. \end{aligned} \quad (4.15)$$

□

4.2 Product of Ginibre matrices with one truncated unitary matrix

This model refers to the product

$$Y = X_M \cdots X_2 V, \quad (4.16)$$

where X_i , $i = 2, \dots, M$ is a Ginibre matrix of size $(n + \nu_i) \times (n + \nu_{i-1})$ with $\nu_i \geq 0$. The $(n + \nu_1) \times n$ matrix V is the left upper block of an $l \times l$ Haar distributed unitary matrix U with $l \geq 2n + \nu_1$. It is known that the squared singular values of V are distributed on $(0, 1)$ according to a Jacobi unitary ensemble; cf. [32].

By [37, Proposition 4.4], we have that the squared singular values of Y form a determinantal process with the correlation kernel

$$\begin{aligned} K_n(x, y) &= \frac{1}{(2\pi i)^2} \int_{-1/2-i\infty}^{-1/2+i\infty} ds \oint_{\Sigma} dt \\ &\quad \prod_{j=0}^M \frac{\Gamma(s + \nu_j + 1)}{\Gamma(t + \nu_j + 1)} \frac{\Gamma(t - n + 1)}{\Gamma(s - n + 1)} \frac{\Gamma(t + n + \kappa)}{\Gamma(s + n + \kappa)} \frac{x^t y^{-s-1}}{s - t}, \end{aligned} \quad (4.17)$$

where

$$\nu_0 = 0, \quad \kappa := l + 1 - 2n > \nu_1, \quad (4.18)$$

Σ is a closed contour going around $0, 1, \dots, n-1$ in the positive direction and $\text{Re } t > -1/2$ for $t \in \Sigma$. To state the universal results for the correlation kernel, we need the following parametrization

$$x_0 = \frac{\left(\sin\left(\frac{M+1}{2}\varphi\right)\right)^{\frac{M+1}{2}}}{\sin \varphi \left(\sin\left(\frac{M-1}{2}\varphi\right)\right)^{\frac{M-1}{2}}}, \quad 0 < \varphi < \frac{2\pi}{M+1}, \quad (4.19)$$

which is a one-to-one mapping from $(0, 2\pi/(M+1))$ to $\left(0, (M+1)^{\frac{M+1}{2}} / \left(2(M-1)^{\frac{M-1}{2}}\right)\right)$; see [23].

Theorem 4.2 (Bulk universality). *Let $K_n(x, y)$ be the correlation kernel defined in (4.17). For $x_0 \in \left(0, (M+1)^{\frac{M+1}{2}} / \left(2(M-1)^{\frac{M-1}{2}}\right)\right)$, which is parametrized by $\varphi \in (0, 2\pi/(M+1))$ through (4.19), we have, with $M \geq 2$ and $\nu_1, \dots, \nu_M, \kappa$ being fixed,*

$$\lim_{n \rightarrow \infty} \frac{e^{-\pi\xi \cot \frac{\varphi}{2}} n^{M-2}}{e^{-\pi\eta \cot \frac{\varphi}{2}} \rho(\varphi)} K_n \left(n^{M-1} \left(x_0 + \frac{\xi}{n\rho(\varphi)} \right), n^{M-1} \left(x_0 + \frac{\eta}{n\rho(\varphi)} \right) \right) = K_{\sin}(\xi, \eta) \quad (4.20)$$

uniformly for ξ and η in any compact subset of \mathbb{R} , where the function ρ is given by

$$\rho(\varphi) = \frac{1}{\pi x_0} \left(\frac{\sin \left(\frac{M+1}{2} \varphi \right)}{\sin \left(\frac{M-1}{2} \varphi \right)} \right)^{1/2} \sin \frac{\varphi}{2}. \quad (4.21)$$

When the reference point x_0 is taken to be the right ending point, we have

Theorem 4.3 (Soft edge universality). *With the correlation kernel $K_n(x, y)$ defined in (4.17), we have, with $\nu_1, \dots, \nu_M, \kappa$ being fixed,*

$$\lim_{n \rightarrow \infty} \exp \left\{ \left(\frac{n}{2} \right)^{\frac{1}{3}} \frac{(M+1)^{\frac{1}{2}}}{(M-1)^{\frac{1}{6}}} (\eta - \xi) \right\} n^{M-\frac{5}{3}} c_2 K_n \left(n^{M-1} \left(x_* + \frac{c_2 \xi}{n^{\frac{2}{3}}} \right), n^{M-1} \left(x_* + \frac{c_2 \eta}{n^{\frac{2}{3}}} \right) \right) = K_{\text{Ai}}(\xi, \eta) \quad (4.22)$$

uniformly for ξ and η in any compact subset of \mathbb{R} , where

$$x_* = \frac{(M+1)^{\frac{M+1}{2}}}{2(M-1)^{\frac{M-1}{2}}}, \quad \text{and} \quad c_2 = \frac{(M+1)^{\frac{M+1}{2}}}{2^{\frac{4}{3}}(M-1)^{\frac{M}{2}-\frac{7}{6}}}. \quad (4.23)$$

Sketched proofs of Theorems 4.2 and 4.3. We scale the values of x and y in (4.17) such that

$$x = n^{M-1} \left(x_0 + \frac{\xi}{\rho(\varphi)n} \right), \quad y = n^{M-1} \left(x_0 + \frac{\eta}{\rho(\varphi)n} \right) \quad (4.24)$$

in the bulk case or

$$x = n^{M-1} \left(x_* + \frac{c_2 \xi}{n^{\frac{2}{3}}} \right), \quad y = n^{M-1} \left(x_* + \frac{c_2 \eta}{n^{\frac{2}{3}}} \right) \quad (4.25)$$

in the soft edge case, where $\xi, \eta \in \mathbb{R}$. By using Stirling's formula for gamma functions and the reflection formula, in the bulk case it follows that, for n large

$$K_n(x, y) \sim \frac{n^{-M+1}}{(2\pi i)^2} \int_{\mathcal{C}} ds \oint_{\Sigma} dt \frac{e^{n(\hat{F}(ns; x_0) - \hat{F}(nt; x_0))}}{s - t} \left(1 + \frac{\xi}{nx_0\rho} \right)^t \left(1 + \frac{\eta}{nx_0\rho} \right)^{-s} \left(x_0 + \frac{\eta}{\rho n} \right)^{-1} \exp \left\{ \sum_{j=0}^M \left(\nu_j + \frac{1}{2} \right) \log \frac{s}{t} - \left(\kappa - \frac{1}{2} \right) \log \frac{n+s}{n+t} - \frac{1}{2} \log \frac{s-n}{t-n} \right\}, \quad (4.26)$$

where the contours \mathcal{C} and Σ depend on x_0 and

$$\hat{F}(z; x_0) = (M+1)(z \log z - 1) - (1+z)(\log(1+z) - 1) - (z-1)(\log(z-1) - 1) - z \log x_0. \quad (4.27)$$

Since

$$\hat{F}_z(z; x_0) = (M+1) \log z - \log(1+z) - \log(z-1) - \log x_0, \quad (4.28)$$

the saddle point of F satisfies the algebraic equation

$$z^{M+1} + x_0(1 - z^2) = 0. \quad (4.29)$$

One can find two explicit solutions of this equation with the help of (4.19), which are given by

$$w_{\pm} = \left(\frac{\sin(\frac{M+1}{2}\varphi)}{\sin(\frac{M-1}{2}\varphi)} \right)^{1/2} e^{\pm i\frac{\varphi}{2}}; \quad (4.30)$$

cf. [23]. The contours \mathcal{C} and Σ are chosen to be

$$\mathcal{C} = \left\{ z \mid \operatorname{Re} \frac{z}{n} = \operatorname{Re} z_{\pm} = \left(\frac{\sin(\frac{M+1}{2}\varphi)}{\sin(\frac{M-1}{2}\varphi)} \right)^{1/2} \cos \frac{\varphi}{2} \right\}, \quad (4.31)$$

and the deformation of

$$\tilde{\Sigma} = \left\{ z = \left(\frac{\sin(\frac{M+1}{2}\phi)}{\sin(\frac{M-1}{2}\phi)} \right)^{1/2} e^{i\frac{\phi}{2}} \mid -\frac{2\pi}{M+1} \leq \phi \leq \frac{2\pi}{M+1} \right\}, \quad (4.32)$$

in manners similar to the construction of Σ based on $\tilde{\Sigma}$ that we described in Sections 3.1 and 2.2.

We then have $\operatorname{Re} \hat{F}(z; x_0)$ defined in (4.27) attains its global maximum at $z = w_{\pm}$ for $nz \in \mathcal{C}$ and its global minimum at $z = w_{\pm}$ for $z \in \tilde{\Sigma}$. Thus, if $x_0 \in \left(0, (M+1)^{\frac{M+1}{2}} / \left(2(M-1)^{\frac{M-1}{2}}\right)\right)$, like (4.15),

$$\begin{aligned} K_n(x, y) &\sim \frac{n^{-M+1}}{2\pi i x_0} \int_{nw_-}^{nw_+} ds \left(1 + \frac{\xi}{x_0 \rho n}\right)^s \left(1 + \frac{\eta}{x_0 \rho n}\right)^{-s} \\ &\sim n^{-M+2} \rho e^{\pi(\xi-\eta) \cot \frac{\varphi}{2}} \frac{\sin \pi(\xi - \eta)}{\pi(\xi - \eta)}, \end{aligned} \quad (4.33)$$

which is (4.20).

As $x_0 \rightarrow x_*$, we have $z_+ = z_- := z_0 = \sqrt{(M+1)/(M-1)}$. In this case, the integration over the contours around z_0 contributes the most. Note that

$$\hat{F}(z; x_*) = \hat{F}(z_0; x_*) + \frac{(M-1)^2}{6}(z - z_0)^3 + \cdots, \quad z \rightarrow z_0. \quad (4.34)$$

With the formula

$$\begin{aligned} K_n(x, y) &\sim \frac{n^{-M+1}}{(2\pi i)^2} \int_{\mathcal{C}} ds \oint_{\Sigma} dt \frac{e^{n(\hat{F}(ns; x_0) - \hat{F}(nt; x_0))}}{s - t} \left(1 + \frac{c_2 \xi}{n^{\frac{2}{3}} x_*}\right)^t \left(1 + \frac{c_2 \eta}{n^{\frac{2}{3}} x_*}\right)^{-s} \\ &\quad \left(x_* + \frac{c_2 \eta}{n^{\frac{2}{3}}}\right)^{-1} \exp \left\{ \sum_{j=0}^M \left(\nu_j + \frac{1}{2}\right) \log \frac{s}{t} - \left(\kappa - \frac{1}{2}\right) \log \frac{n+s}{n+t} - \frac{1}{2} \log \frac{s-n}{t-n} \right\}, \end{aligned} \quad (4.35)$$

by the change of variables

$$s = nz_0 + n^{\frac{2}{3}}((M-1)^2/2)^{-1/3}u, \quad t = nz_0 + n^{\frac{2}{3}}((M-1)^2/2)^{-1/3}v, \quad (4.36)$$

and by the deformation of \mathcal{C} and Σ such that they go through the vicinity of nz_0 in proper directions, we have

$$K_n(x, y) \sim n^{-M+\frac{5}{3}} c_2^{-1} \left(1 + \frac{c_2 \xi}{x_* n^{\frac{2}{3}}}\right)^{nz_0} \left(1 + \frac{c_2 \eta}{x_* n^{\frac{2}{3}}}\right)^{-nz_0} K_{\text{Ai}}(\xi, \eta). \quad (4.37)$$

Thus (4.22) is proved. \square

Remark 2. By setting $\xi = \eta = 0$ in (4.5) and (4.20), the bulk limit also implies point-wise convergence of one-point correlation functions in the support of the limiting measure. The functions $\rho(\varphi)$'s in (4.5) and (4.20) are actually density functions of the limiting spectral distribution for the squared singular values under proper parametrizations. Similar result holds for products of Ginibre matrices; see Theorem 1.1. Thus we recover the limiting mean density results of the random matrix models discussed above, which were previously derived by the moment method of Stieltjes transforms; see e.g. [8], [22]. However, the moment method has the advantage in the discovery of natural parametrizations like (1.13), (4.4) and (4.19) by combinatorial relations; see [14], [23], [29], [41].

Acknowledgment

We thank the anonymous referees for their careful reading and constructive suggestions. The work of D.-Z. Liu was supported by the National Natural Science Foundation of China (Grants 11301499 and 11171005), and by the Fundamental Research Funds for the Central Universities (Grant WK0010000048). The work of D. Wang was partially supported by the start-up grant R-146-000-164-133. The work of L. Zhang was partially supported by The Program for Professor of Special Appointment (Eastern Scholar) at Shanghai Institutions of Higher Learning (No. SHH1411007) and by Grant SGST 12DZ 2272800, EZH1411513 from Fudan University.

References

- [1] K. Adhikari, N. K. Reddy, T. R. Reddy, and K. Saha. Determinantal point processes in the plane from products of random matrices, 2013. arXiv:1308.6817, to appear in Ann. Inst. Henri Poincaré Probab. Stat.
- [2] M. Adler, P. van Moerbeke, and D. Wang. Random matrix minor processes related to percolation theory. *Random Matrices Theory Appl.*, 2(4):1350008, 72, 2013.
- [3] G. Akemann and Z. Burda. Universal microscopic correlation functions for products of independent Ginibre matrices. *J. Phys. A*, 45(46):465201, 18, 2012.
- [4] G. Akemann, Z. Burda, M. Kieburg, and T. Nagao. Universal microscopic correlation functions for products of truncated unitary matrices. *J. Phys. A*, 47(25):255202, 26, 2014.
- [5] G. Akemann and J. R. Ipsen. Recent exact and asymptotic results for products of independent random matrices, 2015. arXiv:1502.01667.
- [6] G. Akemann, J. R. Ipsen, and M. Kieburg. Products of rectangular random matrices: Singular values and progressive scattering. *Phys. Rev. E*, 88(5):052118, 13, 2013.
- [7] G. Akemann, M. Kieburg, and L. Wei. Singular value correlation functions for products of Wishart random matrices. *J. Phys. A*, 46(27):275205, 22, 2013.
- [8] N. Alexeev, F. Götze, and A. Tikhomirov. Asymptotic distribution of singular values of powers of random matrices. *Lith. Math. J.*, 50(2):121–132, 2010.
- [9] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An introduction to random matrices*, volume 118 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010.

- [10] T. Banica, S. T. Belinschi, M. Capitaine, and B. Collins. Free Bessel laws. *Canad. J. Math.*, 63(1):3–37, 2011.
- [11] R. Beals and J. Szmigielski. Meijer G -functions: a gentle introduction. *Notices Amer. Math. Soc.*, 60(7):866–872, 2013.
- [12] M. Bertola and T. Bothner. Universality conjecture and results for a model of several coupled positive-definite matrices. *Comm. Math. Phys.*, 337(3):1077–1141, 2015.
- [13] M. Bertola, M. Gekhtman, and J. Szmigielski. Cauchy-Laguerre two-matrix model and the Meijer- G random point field. *Comm. Math. Phys.*, 326(1):111–144, 2014.
- [14] P. Biane. Processes with free increments. *Math. Z.*, 227(1):143–174, 1998.
- [15] A. Borodin. Biorthogonal ensembles. *Nuclear Phys. B*, 536(3):704–732, 1999.
- [16] P. Bougerol and J. Lacroix. *Products of random matrices with applications to Schrödinger operators*, volume 8 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1985.
- [17] Z. Burda, R. A. Janik, and B. Waclaw. Spectrum of the product of independent random Gaussian matrices. *Phys. Rev. E (3)*, 81(4):041132, 12, 2010.
- [18] Z. Burda, A. Jarosz, G. Livan, M. A. Nowak, and A. Swiech. Eigenvalues and singular values of products of rectangular Gaussian random matrices. *Phys. Rev. E (3)*, 82(6):061114, 10, 2010.
- [19] T. Claeys, A. B. J. Kuijlaars, and D. Wang. Correlation kernels for sums and products of random matrices, 2015. arXiv:1505.00610.
- [20] A. Crisanti, G. Paladin, and A. Vulpiani. *Products of random matrices in statistical physics*, volume 104 of *Springer Series in Solid-State Sciences*. Springer-Verlag, Berlin, 1993. With a foreword by Giorgio Parisi.
- [21] P. J. Forrester. *Log-gases and random matrices*, volume 34 of *London Mathematical Society Monographs Series*. Princeton University Press, Princeton, NJ, 2010.
- [22] P. J. Forrester. Eigenvalue statistics for product complex Wishart matrices. *J. Phys. A*, 47(34):345202, 22, 2014.
- [23] P. J. Forrester and D.-Z. Liu. Raney distributions and random matrix theory. *J. Stat. Phys.*, 158(5):1051–1082, 2015.
- [24] P. J. Forrester and D.-Z. Liu. Singular values for products of complex Ginibre matrices with a source: hard edge limit and phase transition, 2015. arXiv:1503.07955.
- [25] P. J. Forrester and D. Wang. Muttalib–Borodin ensembles in random matrix theory — realisations and correlation functions, 2015. arXiv:1502.07147.
- [26] H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Statist.*, 31:457–469, 1960.
- [27] F. Götze, H. Kösters, and A. Tikhomirov. Asymptotic spectra of matrix-valued functions of independent random matrices and free probability, 2014. arXiv:1408.1732, to appear in *Random Matrices Theory Appl.*

- [28] F. Götze and A. Tikhomirov. On the asymptotic spectrum of products of independent random matrices, 2010. arXiv:1012.2710.
- [29] U. Haagerup and S. Möller. The law of large numbers for the free multiplicative convolution. In *Operator algebra and dynamics*, volume 58 of *Springer Proc. Math. Stat.*, pages 157–186. Springer, Heidelberg, 2013.
- [30] J. R. Ipsen and M. Kieburg. Weak commutation relations and eigenvalue statistics for products of rectangular random matrices. *Phys. Rev. E*, 89(3):032106, 20, 2014.
- [31] M. E. H. Ismail. *Classical and quantum orthogonal polynomials in one variable*, volume 98 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2009. With two chapters by Walter Van Assche, With a foreword by Richard A. Askey, Reprint of the 2005 original.
- [32] T. Jiang. Approximation of Haar distributed matrices and limiting distributions of eigenvalues of Jacobi ensembles. *Probab. Theory Related Fields*, 144(1-2):221–246, 2009.
- [33] K. Johansson. Random matrices and determinantal processes. In *Mathematical statistical physics*, pages 1–55. Elsevier B. V., Amsterdam, 2006.
- [34] R. P. Kanwal. *Linear integral equations*. Birkhäuser Boston, Inc., Boston, MA, second edition, 1997.
- [35] M. R. Krishnapur. *Zeros of random analytic functions*. ProQuest LLC, Ann Arbor, MI, 2006. Thesis (Ph.D.)—University of California, Berkeley.
- [36] A. B. J. Kuijlaars. Universality. In *The Oxford handbook of random matrix theory*, pages 103–134. Oxford Univ. Press, Oxford, 2011.
- [37] A. B. J. Kuijlaars and D. Stivigny. Singular values of products of random matrices and polynomial ensembles. *Random Matrices Theory Appl.*, 3(3):1450011, 22, 2014.
- [38] A. B. J. Kuijlaars and L. Zhang. Singular values of products of Ginibre random matrices, multiple orthogonal polynomials and hard edge scaling limits. *Comm. Math. Phys.*, 332(2):759–781, 2014.
- [39] D.-Z. Liu, C. Song, and Z.-D. Wang. On explicit probability densities associated with Fuss-Catalan numbers. *Proc. Amer. Math. Soc.*, 139(10):3735–3738, 2011.
- [40] Y. L. Luke. *The special functions and their approximations, Vol. I*. Mathematics in Science and Engineering, Vol. 53. Academic Press, New York-London, 1969.
- [41] T. Neuschel. Plancherel-Rotach formulae for average characteristic polynomials of products of Ginibre random matrices and the Fuss-Catalan distribution. *Random Matrices Theory Appl.*, 3(1):1450003, 18, 2014.
- [42] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*, volume 335 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, Cambridge, 2006.
- [43] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, editors. *NIST handbook of mathematical functions*. U.S. Department of Commerce, National Institute of Standards and Technology, Washington, DC; Cambridge University Press, Cambridge, 2010. With 1 CD-ROM (Windows, Macintosh and UNIX).

- [44] S. O'Rourke and A. Soshnikov. Products of independent non-Hermitian random matrices. *Electron. J. Probab.*, 16:no. 81, 2219–2245, 2011.
- [45] K. A. Penson and K. Życzkowski. Product of Ginibre matrices: Fuss-Catalan and Raney distributions. *Phys. Rev. E*, 83(6):061118, 9, 2011.
- [46] A. Soshnikov. Determinantal random point fields. *Uspekhi Mat. Nauk*, 55(5(335)):107–160, 2000.
- [47] C. A. Tracy and H. Widom. Level spacing distributions and the Bessel kernel. *Comm. Math. Phys.*, 161(2):289–309, 1994.
- [48] A. M. Tulino and S. Verdú. Random matrix theory and wireless communications. *Found. Trends Commun. Inform. Theory*, 1(1):1–182, 2004.
- [49] L. Zhang. A note on the limiting mean distribution of singular values for products of two Wishart random matrices. *J. Math. Phys.*, 54(8):083303, 8, 2013.
- [50] L. Zhang. Local universality in biorthogonal Laguerre ensembles, 2015. arXiv:1502.03160.